Bayesian analysis of fluorescence lifetime imaging data

Rowley, Mark

Awarding institution:
King's College London

Title: Bayesian analysis of fluorescence lifetime imaging data

Author: Mark Rowley

BAYESIAN ANALYSIS OF FLUORESCENCE LIFETIME IMAGING DATA

MARK IAN ROWLEY

2013

THIS THESIS IS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF PHD IN KING'S COLLEGE LONDON

I wish to thank the EPSRC/BBSRC (EP/C546105/1 and EP/C546113/1) for financial support, and my supervisors B Vojnovic, T Coolen and P Barber for the opportunity to work in this subject area and for many useful discussions. I am also indebted to many people who make FLIM experiments in cells and tissue possible, including but not limited to, the biologists, T Ng, G Weitsman, at the Randall and Cancer Divisions, KCL, and the instrument builders, I Tullis, R Newman and J Prentice at the Gray Institute in Oxford.

Abstract

The development of a novel photon-by-photon Bayesian analysis for time-domain Fluorescence Lifetime Imaging Microscopy (FLIM) data, and its application to both real experimental biological and synthetic data, is presented in this thesis. FLIM is an intensity-independent and sensitive optical technique for studying the cellular environment and can robustly exploit Förster Resonance Energy Transfer (FRET) to enable protein-protein interactions to be located within living or fixed cells. Careful analysis of fluorescence lifetime data, often comprising multi-exponential kinetics, is crucial to elucidating FRET via FLIM.

The developed Bayesian analysis is demonstrated to offer more accurate fitting of data with lower photon counts, allowing greater acquisition speeds. As well as revealing information previously unobtainable, such as direct error estimates, fitting model probabilities, and instrument response extraction, the developed approach allows for future extensions which can exploit the full probability distribution. In a section of this work already published [1], Bayesian mono-exponential analysis was shown to offer robust estimation with greater precision at low total photon counts, estimating fluorescent lifetimes to a level of accuracy not obtained using other techniques. Bayesian mono-exponential parameter estimates obtained with the developed Bayesian analysis are improved compared to those obtained using maximum likelihood, least squares, and the phasor data fitting approaches. In this work, Bayesian bi-exponential analysis based on an improved fully-analytic time-domain FLIM system model is shown to also offer improved decay parameter estimates.

The developed analysis offers fluorescence decay model selection by exploiting the hierarchical nature of Bayesian analysis. This innovation enables the quantitative determination of the likelihood of the data being due to mono- or bi-exponential decay processes, for example. Model selection applied to FLIM promises to simplify processing where the exact kinetics are not known. Finally, the determination of an approximated instrument response function from observed fluorescence decay data alone is also possible.

Contents

Part I

Introduction

This thesis presents a Bayesian analysis for fluorescence lifetime imaging microscopy (FLIM) data. FLIM is widely[1] used in the biomedical sciences, and has been used extensively for the study of protein-protein interactions and the molecular environment in the laboratories in the Randall Division of Cell & Molecular Biophysics at King's College London and at the Gray Institute for Radiation Oncology and Biology at the University of Oxford, where most of the work that has culminated in this thesis has been conducted. The primary aim of the research was to develop a robust analysis technique for time-domain FLIM data targeted particularly towards those situations when data is in short supply and its analysis is not straightforward. The provision of a reliable technique can thereby reduce the time taken to acquire such data and potentially allow a greater insight into dynamic processes within cells and also reduce experiment time.

Although FLIM has found a niche in the life sciences for the study of protein-protein interaction e.g. [2, 3], it has been used in a diverse range of studies outside of this area, having found application in combustion research to study turbulent hot nitrogen flows [4], in microfluidic system research to observe dynamic fluid mixing in real-time [5], for refractive index sensing [6], and in viscosity measurements by using fluorescent molecular rotors [7] in molecules and proteins. In the past decade FLIM has also been used in a few less conventional settings, from a demonstration of its potential for the analysis of gunshot residue in forensic science [8], to its use as a tool for the non-destructive analysis of Renaissance fresco paintings in Italy [9] and the fluorescence lifetime mapping of Michelangelo's David [10]. Indeed, even the humble tomato has been able to avoid neither the advances of the wilt virus nucleosapid protein nor FLIM [11, 12].

Before embarking on the main subject matter of this thesis, time-resolved FLIM data and its analysis, a concise introduction to fluorescence and some of the key technologies that are essential to enabling FLIM is provided. Fluorescent proteins, their introduction into the cellular environment as fluorescent probes, and the fundamental limits that restrict the rate of emission and detection of fluorescence photons in FLIM, are discussed. A brief history of the technical development of fluorescence microscopy is given in [13].

Fluorescence [14, 15] is one of a family of processes described by the phenomenon of luminescence. Certain molecules emit light from electronically excited states generated

[1] As of March 2013, a search for publications including "FLIM" or "fluorescence lifetime" in the public biomedical literature reference database "PubMed" (http://www.ncbi.nlm.nih.gov/pubmed) yields nearly 6000 publications, the number of publications per year having increased year-on-year in sixteen of the past twenty years, and with 492 publications in 2012 alone.

by e.g. physical (particularly absorption of optical or higher energy photons), mechanical (e.g. friction), or chemical mechanisms. The generation of luminescence through excitation of a molecule by light is a phenomenon termed photoluminescence, which is usually divided into two categories, fluorescence and phosphorescence, depending on the electronic configuration of the excited state and the specific emission pathway associated with a molecules of specific substances. Fluorescence is the property of some atoms and molecules to absorb light at a particular wavelength and to subsequently emit light at a longer wavelength following a very short interval: the fluorescence lifetime. The process of phosphorescence occurs in a similar way to fluorescence, but is associated with a much longer excited state lifetime. The fluorescence process is determined by three important events, all of which occur on timescales that are separated by several orders of magnitude. Excitation of an appropriate molecule by an incoming photon happens in femtoseconds, while vibrational relaxation of excited state molecule electrons to the lowest energy level is much slower and is usually over in picosecond timescale. The final process, emission of a less energetic, longer wavelength photon, compared to the excitation photon. The return of the molecule to the ground state, occurs in nanoseconds. Other processes such as delayed fluorescence and phosphorescence take much longer to return to the ground state. They are included here for completeness, but are not exploited in the work presented here. A simplified Jablonski diagram [16, 17] illustrating the energy transition processes occurring during luminescence is shown in Fig. 1.

An illustration of the shifting of the fluorescence emission spectra to longer, less energetic, wavelengths relative to the absorption spectra is shown in Fig. 2 (a). The difference between the peak excitation (absorption) and emission wavelengths is known as the Stokes Shift [14, 15], as illustrated in Fig. 2 (a). In the case that there is spectral overlap between the emission spectra of a fluorophore and the excitation spectra of a chromophore, as shown in Fig. 2 (b), energy may be transferred non-radiatively between the two by means of resonance energy transfer, as discussed below.

Förster (or fluorescence) resonance energy transfer (FRET) [18, 19] is the process by which a fluorophore (the donor) in an excited state transfers energy non-radiatively to a nearby chromophore (the acceptor). FRET occurs at donor-acceptor separations of between about 1 nm and 10 nm [20], and is therefore very well suited to studying biological systems and has been described as a "spectroscopic ruler" [21]. The occurrence of resonance energy transfer requires that there be overlap between the emission spectra of a

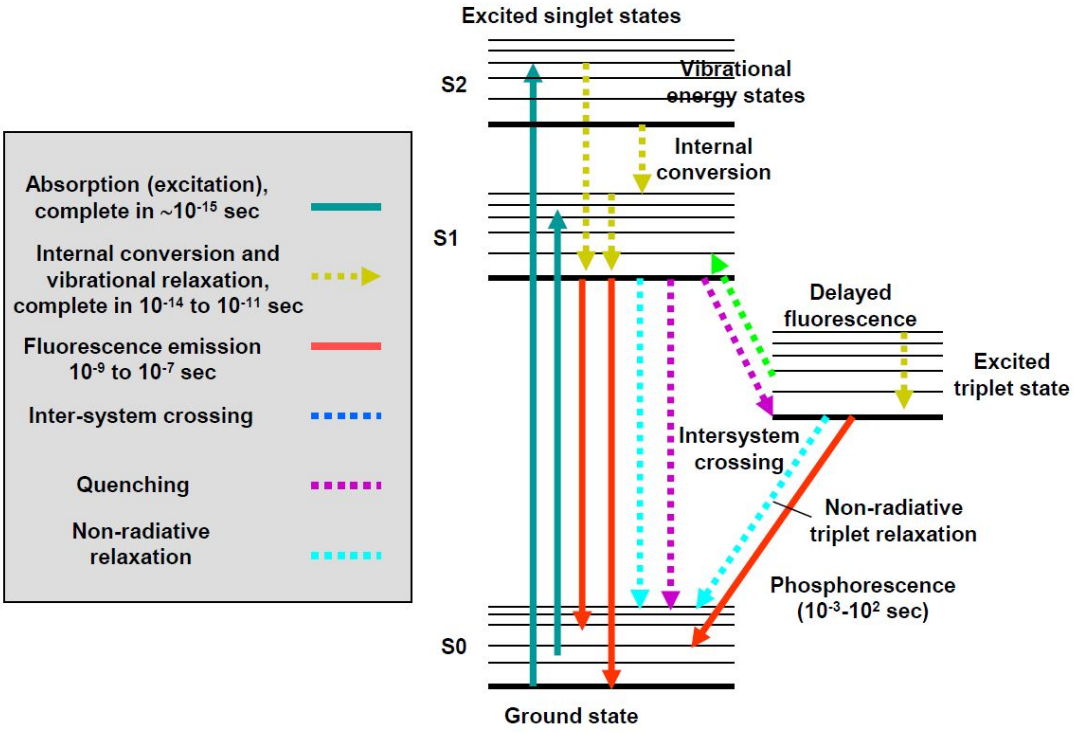
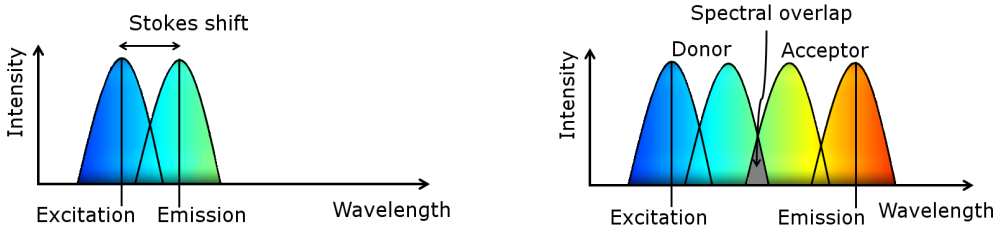


Figure 1: *Energy transition processes occurring during luminescence, depicted as a simplified Jablonski diagram:* A Jablonski diagram illustrating the processes present in fluorescence, that is, excitation from the ground state (S0), internal vibrational relaxation to the first excited state (S1), and the emission of a fluorescence photon for a return to the ground state (S0). For completeness, the phosphorescence and delayed fluorescence pathways are also shown.

donor fluorophore and the excitation spectra of an acceptor chromophore, as illustrated in Fig. 2. The FRET process is illustrated in Fig. 3 by a simplified Jablonski diagram.

In Fig. 4, the use of FRET to study protein-protein interactions is illustrated. If two proteins, for example, one labeled with blue fluorescent protein (BFP) (the donor) and the other with green fluorescent protein (GFP) (the acceptor), physically interact, then increased intensity at the acceptor emission maximum (510 nanometers) will be observed when the complex is excited at the maximum absorbance wavelength (380 nanometers) of the donor. Failure of the proteins to form a complex results in no acceptor (GFP) fluorescence emission. An illustration of how intramolecular FRET can be used to detect protein conformational change is shown in Fig. 5; by quantifying the FRET efficiency



(a) Stokes Shift

(b) FRET spectral overlap

Figure 2: *Fluorescence emission and absorption spectra:* In (a) an illustration of the Stoke's Shift, the difference between the peak excitation (absorption) and emission wavelengths [14, 15], and in (b) an illustration of the spectral overlap between the emission spectra of a donor fluorophore and the excitation spectra of an acceptor chromophore as required for resonance energy transfer.

experimentally and applying Eqn. (1.1) the distance between the fluorochromes can be determined and offer insight as to the structure of the protein.

In order to utilise fluorescence and FRET for microscopy it is necessary, therefore, to excite a fluorescent substance in a sample and to detect the fluorescence emission. It is also usually necessary to somehow introduce a fluorescent substance(s) into the sample, as discussed below.

Since the development of immunofluorescence e.g. [22, 23], which enabled the distribution of specific molecules within a cell, having been tagged using antibodies labelled with fluorescent dyes, to be visualised using fluorescence microscopy, the range of fluorophores available to probe the the cellular environment has increased significantly. The development of optimised fluorescent proteins as tags, which can be fused to a wide variety of targets, has been key to making fluorescence microscopy the widely used tool that it is in the biomedical sciences [24]. Using such fluorescent probes and fluorescence microscopy techniques has enabled biological processes to be studied with unprecedented spatial and temporal precision [25]. The green fluorescent protein[2] (GFP) [26, 27] and its variants, for example, can be fused to a wide variety of targets [24]. A cartoon illustration of GFP is shown in Fig. 6, as is a picture of mice expressing enhanced GFP under UV-

[2]In 2008, The Nobel Prize for Chemistry was awarded to Osamu Shimomura, Martin Chalfie, and Roger Y. Tsien, "for the discovery and development of the green fluorescent protein, GFP". Having been first discovered in the jellyfish Aequorea victoria [26], GFP is widely exploited for research in biochemistry and cell biology [27].

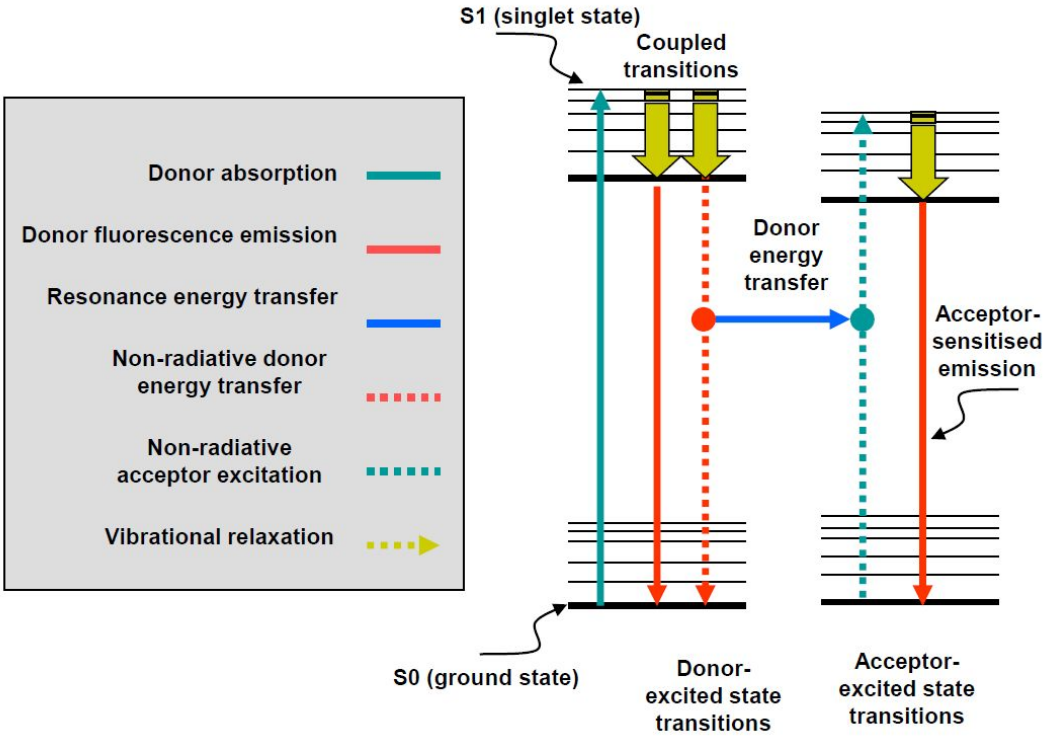


Figure 3: *FRET, depicted as a simplified Jablonski diagram:* Jablonski diagram illustrating coupled transitions involved between donor emission and acceptor absorption occurring during Frster resonance energy transfer. For simplicity, higher vibrational states present during donor excitation have been omitted.

illumination [28]. The interested reader is directed to e.g. [24, 29, 30, 25, 31] for reviews of fluorescent probes, their benefits and limitations for imaging in cell biology, and their importance to fluorescence microscopy-based research in the biomedical sciences. The development of FRET biosensor probes, and their use for the study of protein-protein interactions, protein conformational changes, and the concentrations of small molecules are reviewed in e.g. [32, 33, 34].

Autofluorescence describes the fluorescent emission due to endogenous fluorophores within a substance, and is often an unwanted source of interference in FLIM experiments [35]. Such natural fluorescence is present in most organisms, and has been characterised, amongst others, in rat hepatocytes, Drosophila melanogaster, numerous fungi, plants such as maize, and alfalfa [35]. Autofluorescence has also been employed to the advantage of

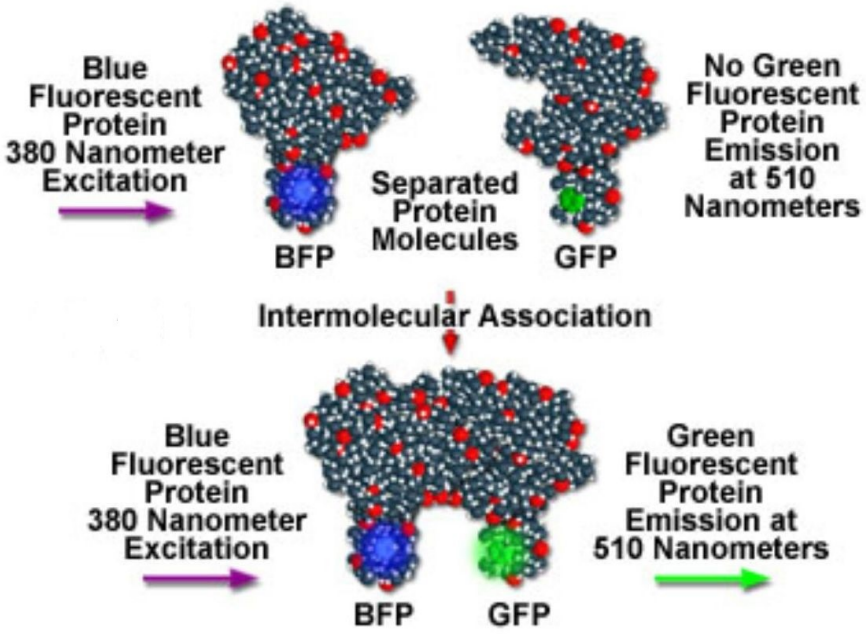


Figure 4: *Intermolecular FRET:* An illustration of protein-protein interaction determination using FRET. If two proteins, one labeled with blue fluorescent protein (BFP) (the donor) and the other with green fluorescent protein (GFP) (the acceptor), physically interact, then increased intensity at the acceptor emission maximum (510 nanometers) will be observed when the complex is excited at the maximum absorbance wavelength (380 nanometers) of the donor. Failure of the proteins to form a complex results in no acceptor (GFP) fluorescence emission. Figure from http://micro.magnet.fsu.edu/primer/techniques/fluorescence/fret/fretintro.html with permission.

basic research and diagnostics [36], for example, to study renal function [37], and as a method for toxicity testing in cells [38].

The detection of the fluorescence emission from a sample is now discussed, hinting at why a robust FLIM data analysis technique targeted particularly towards those situations when data is in short supply, such as is introduced in this thesis, may be required.

The efficiency with which a particular fluorophore absorbs a photon of the excitation light is a function of the molecular cross-section, and the likelihood of photon absorption taking place is known as the *extinction coefficient*. Larger extinction coefficients indicate that it is more likely that a photon of appropriate wavelength region is more likely to be absorbed. Absorption is linearly proportional to the excitation intensity for low fluorophore

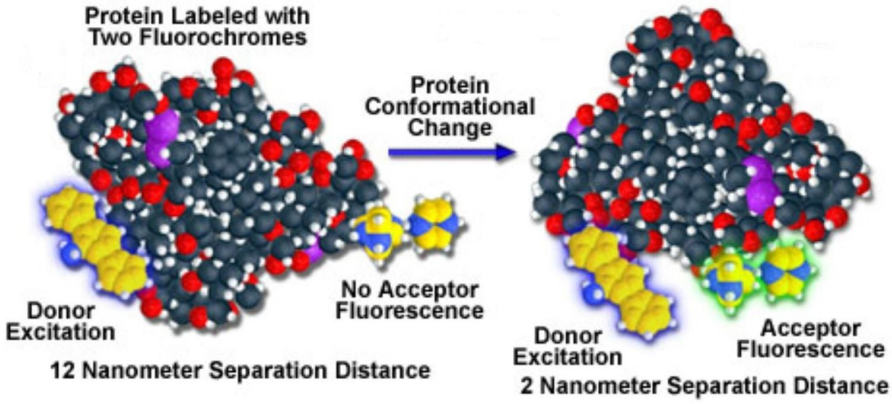


Figure 5: *Intramolecular FRET:* An illustration of intramolecular FRET for the detection of protein conformational change, the donor fluorophore and acceptor chromophore being sited at opposite ends of the same protein. In the native state, on the left, the two fluorophores are sufficiently separated such that FRET does no occur, and, on the right, with the protein in a different confirmation and the donor and acceptor in close proximity such that resonance energy transfer occurs. The protein conformational change can therefore inferred by the detection of FRET. Figure from http://micro.magnet.fsu.edu/primer/techniques/fluorescence/fret/fretintro.html with permission.

concentrations and the average time during which the electric dipole of the fluorophore is aligned with the excitation electromagnetic waves electric vector (polarisation).

The *quantum yield* of the fluorophore describes the ratio of the number of photons emitted compared to those that are absorbed. Since, by definition, fluorophores emit some light when they are excited, the quantum yield must have a value that exceeds zero and cannot be higher than unity. Quantum yield values below 1 result from the loss of energy through non-radiative pathways, (for example through heat or through energy required for photo-chemical reactions). High quantum yields indicate a higher probability that re-radiation, through fluorescence, will take place. In practice, fluorophores with quantum yields in the range of 0.2-0.99 are used in fluorescence microscopy. The fluorophore extinction coefficient, its quantum yield, its fluorescence lifetime and the intensity of the excitation source are factors of crucial importance that determine the utility of a particular fluorophore. Fundamental limits within the process of fluorescence provide limiting values for these quantities [39, 40]. Fluorescence lifetimes represent a statistical overview of a stochastic process, variations within which are slow compared to the rate of illumi-

(a) GFP (b) 'GFP mice'

Figure 6: *Green Fluorescent Protein (GFP):* In (a) a GFP model (left) and a picture of a Aequorea victoria jellyfish [26, 27], and in (b) two mice expressing enhanced GFP under UV-illumination flanking one plain mouse [28].

nation (photon flux). This results in a random delay with respect to the emission (with respect to excitation) and no phase information from the illumination can be recovered from fluorescence. Regardless of the light absorbed (e.g. laser light with high spatial and temporal coherence) fluorescence is always temporally incoherent. The total fluorescence output F can be determined from the product of cross-section, quantum yield and excitation intensity:

$$F = \sigma Q I \tag{1}$$

where σ is the molecular absorption cross-section, Q is its quantum yield, and I is the excitation light flux, typically $2 - 3 \times 10^{20}$ photons sec^{-1} cm^{-2}. A 'typical' fluorophore has an absorption cross-section of $\sim\ 2 \times 10^{-16}$ cm^2 molecule^{-1}, and a quantum yield of 0.8, resulting in a value for F of 3.2×10^4 photons per second per molecule.

A typical fluorescent molecule can emit about 10^4 photons before destructive quenching takes place. Free radicals, generated as a by-product of excitation are responsible for this destruction. Thus, the signal can be collected, using maximum illumination, for a few seconds before the probe is destroyed. The collection time is longer when the illumination flux is decreased, as it generally is in practice.

The time constants associated with absorption, relaxation and emission mean that significant fluorescence from a population of molecules may last up to around 10 ns, though

this is often somewhat shorter (e.g. the fluorophores used in this thesis have lifetimes of ~ 2 ns). If we consider a maximum time constant of 10 ns, cycling between the ground and excited states occurs at 10^8 photons per second. Any attempt to increase this is likely to lead to saturation (i.e. the generation of a population inversion where more molecules are in the excited state than are in the ground state). Imaging systems which rely on excitation beam scanning often operate at excitation intensities close to saturation, though in the instrument described in Section 1.2.1, this is not the case, as the usable emission photon rates are limited by the counting rate of the detection system used.

Whether or not the fluorescence emission, in a given spectral emission band, can yield a useful signal depends on the ability of the objective to collect the emission, on the transmission characteristics of the optical system used, and on the quantum efficiency and noise of the detector.

The concentration of fluorophores in a cell is now considered. When a high concentration is present, the fluorophores will self-quench through inter-molecular reactions, which remove energy from the excited state without emission of fluorescence. A typical value below which such self-quenching does not occur is 10^{-6} M [39, 40] and therefore the upper limit to molecule number is $\sim 10^6$ molecules per cell, where a typical cell volume of 1 pl is assumed. The limit to the signal available is thus 10^{10} photons per cell, generated by a maximum illumination of 10^{13} photons per second.

When detecting a photon, a statistical uncertainty is always present, irrespective of the detector used, as described by Poisson statistics. When photon numbers are larger (e.g $\gtrsim 15$) the Poisson statistics are approximated by Gaussian statistics and both the noise and the signal-to-noise ratio is equal to the square root of the photon count. The 1 μM probe can thus be detected with a maximum signal-to-noise ratio of 10^5 when a perfect detection system is used. However, when the distribution of fluorescence is to be imaged, and when the fluorescence is to be sampled in time, the signal-to-noise ratio is reduced by dividing by the square root of the number of space and time samples.

Fluorescence emission is isotropic and even the highest numerical aperture objectives (1.4 NA oil immersion) can only collect $\sim 30\%$ of the emitted light. Additional losses in the optics of the instrument further reduce the available photons so that typically no more than 10% of the emission is available for detection. The detector quantum efficiency, DQE, which, in the case of electronic detectors, describes the likelihood of converting a photon to an electron, rarely exceeds 80% (in the detectors used in this thesis, the detec-

tor DQE $\sim 20\%$) further restricts the signal-to-noise ratio. We can thus record no more than $10^8 - 10^9$ counts at no more than $10^{12} - 10^{13}$ counts per second from the whole cell. Imaging the cell across $\sim\ 10^4$ pixels, provides us with a maximum pixel count of $10^4 - 10^5$ at a maximum count rate of $10^8 - 10^9$. In practice, of course lower fluorophore concentrations are often present and the collected counts and the collection rate are consequently reduced. It is this low photon count availability that triggered the work described here: while imaging can be performed with a reasonable signal-to-noise ratio, determining lifetimes is considerably harder, always forcing a compromise between high signal-to-noise ratios with low spatial resolution and *vice-versa*. Any time-dependant changes are clearly even harder to determine. The 'operating window' available to advanced multi-dimensional fluorescence acquisition is narrow, but fortunately accessible with modern instrumentation.

An interesting possible everyday use of FLIM was presented in [41] in which its potential application as an anti-counterfeiting technology for paper money was demonstrated; genuine U.S. Federal Reserve bank notes exhibit an intrinsic two component fluorescence decay "signature" that, it is suggested, differs little from note to note and denomination to denomination, and is only 'minimally influenced by wear-and-tear and/or residual skin oils from regular handling'. Of the tested counterfeit \$100 notes, four of nine exhibited only a mono-exponential fluorescence decay and the remaining five could be determined as not being genuine due to significant deviations in the longer decay component lifetime. In essence, a simple model selection test to discriminate between a mono-exponential decay and a bi-exponential decay (performed in [41] by visual inspection of the fitted decay curve and residuals) was applied to identify the four of nine counterfeit notes in the simple case that the counterfeits exhibited a mono-exponential decay. It is noted in [41] that even should counterfeiters be able to acquire the fluorescence decay "signature" of their counterfeits it would be difficult to tweak it in order to mimic that of genuine notes.

In this work, Bayesian techniques are applied (both in theory and with implemented software) with an intended use of studying the cellular environment through FLIM data. These methods incorporate a detailed mathematical model of the practical arrangement of excitation and detection from an imaging system. In order to aid readability, wherever possible, any of the mathematical and technical intricacies required in this work are detailed in the Appendices rather than the main body of the text.

In the remainder of this introductory part of the thesis, FLIM is introduced in greater detail and the existing analysis methods are discussed. The application of FLIM to protein-protein interaction studies is discussed in Chapter 1, along with a description of the most commonly used FLIM data acquisition methods, and a discussion of some of the practicalities surrounding FLIM data analysis and how the analysis approach may impact the rest of the experimental chain. This is followed in Chapter 2 by a review of the different FLIM data analysis techniques that are currently commonly used for the quantification of FLIM fluorescence decay signals, and the chapter contains an appraisal of their relative advantages and shortcomings.

The theoretical development of the Bayesian analysis for time-domain FLIM is presented in Part II (Chapters 3, 4, & 5), beginning in Chapter 3 with a stand-alone introduction to the Bayesian techniques which are the basis of the research presented here. The introduction of the fully-analytic model for time-domain FLIM which is at the core of this work follows in Chapter 4, and can be read independently or in conjunction with Appendix A, which contains the step-by-step mathematical development of the model. In Chapter 5 the Bayesian methodology is applied to the developed time-domain FLIM system model for the purpose of fluorescence decay parameter estimation, fluorescence decay parameter estimation with simultaneous FLIM instrument parameter estimation, and fluorescence decay model selection.

The results of the application of the developed Bayesian algorithms are presented in Part III (Chapters 6, 7, 8, & 9), in which the performance of the Bayesian algorithms is demonstrated, using both real experimental data and synthetically generated data, and compared to those of the commonly used time-domain FLIM data analysis methods. The application of the Bayesian mono-exponential and bi-exponential fluorescence decay analysis algorithms are discussed in Chapters 6 & 7 respectively. The results of applying the developed Bayesian analysis for the simultaneous estimation of fluorescence decay and FLIM instrument parameters are then presented in Chapter 8, for a number of different (plausible) synthetic instruments and also with the experimental data obtained for [1]. In Chapter 9 the performance of the developed Bayesian fluorescence decay model selection algorithm is discussed.

Part IV reflects upon the developed Bayesian techniques, critically discusses some of the advantages they offer over established FLIM data analysis techniques, and provides some suggestions for future research.

Chapter 1

Fluorescence lifetime imaging microscopy (FLIM)

Fluorescence lifetime imaging microscopy (FLIM) is a powerful optical technique that finds its primary application in the biomedical sciences for the study of molecular environments. The fluorescence lifetime of a fluorophore can depend on a range of biological and chemical properties of interest [42]. Using FLIM to observe the spatial distribution of fluorescence lifetimes of fluorophores within a sample provides a means of observing the behaviour and variation of sample properties [43] such as ion concentrations, hydrophobicity, oxygen concentration, molecular binding, and molecular interaction, by energy transfer when two proteins are close to each other.

In performing FLIM, fluorophores within a sample being studied are optically excited and photons from the resulting fluorescence decay process(es) are detected. The fluorescence lifetime is the characteristic time that a fluorescent molecule remains in the excited state before returning to its ground state, and is estimated (along with other parameters that describe a fluorescent decay) by analysis of the detected photon data. Usually, a FLIM image is formed by collecting FLIM data at many points over the sample being studied [44] (either by laser scanning or otherwise), and the spatial variation of the fluorophores and their lifetimes are mapped.

FLIM benefits from the significant advantage that the lifetime of a fluorescence decay does not depend on fluorophore concentration nor on the excitation light or intensity [45]. FLIM is used in some laboratories, including in the Randall Division of Cell & Molecular

Biophysics where this work has been conducted, in live and fixed cell and tissue imaging studies as it is non-destructive and only minimally invasive [42]. In designing and setting the expectations for a FLIM experiment there are, however, many conflicting ideals which must be traded against each other, according to the sample being studied, the time available, the availability and financial cost of equipment. In using FLIM to study, say, protein-protein interactions the reliable analysis of fluorescence decay photon data for the accurate quantification of an underlying fluorescence decay process is desired. The accuracy of the analysis is improved by acquiring data sets that contain large photon numbers but the acquisition of such data sets requires either an increased imaging duration or a more intense illumination on the sample or both. In imaging a sample for longer, the temporal resolution at which dynamic processes can be studied is sacrificed whereas in increasing the excitation power the sample itself may be sacrificed. It is clear that to benefit from the many advantages FLIM offers, it is necessary to somehow balance the data and accuracy requirements of an analysis with the data availability for the experiment. We will see later why this tradeoff is particularly pertinent to many studies that employ FLIM techniques, especially those in the biomedical sciences.

In the next section, the application of FLIM to Förster (or Fluorescence) Resonance Energy Transfer (FRET) studies is discussed, followed by a discussion of some of the other uses of FLIM. Attention then moves to the acquisition of FLIM data in both the frequency-domain and the time-domain, the two most popular FLIM acquisition methods, without focusing heavily on the technical details. This chapter concludes with a discussion of the need for accurate FLIM data analysis; this serves as an introduction to a review of the past and present literature pertaining to fluorescence decay analysis, leading on to the main subject of this thesis, the analysis of time-domain FLIM data acquired using time correlated single photon counting (TCSPC) [46].

1.1 Application of FLIM to fluorescence resonance energy transfer (FRET)

The use of FLIM to determine Förster (or fluorescence) resonance energy transfer (FRET) [18, 19] is a versatile and sensitive optical technique for studying the cellular environment and enables protein-protein interactions on the nanometer scale within living or fixed cells to be located with a micrometer scale [47, 48]. FRET is the process by which a fluorophore (the donor) in an excited state transfers energy non-radiatively to a nearby chromophore (the acceptor), and can occur at separations of between about 1 nm and 10 nm [20]. The proportion of photons absorbed by the donor that are transferred to the acceptor, the FRET efficiency E, is very sensitive to the separation r between the donor and the acceptor, and is given by,

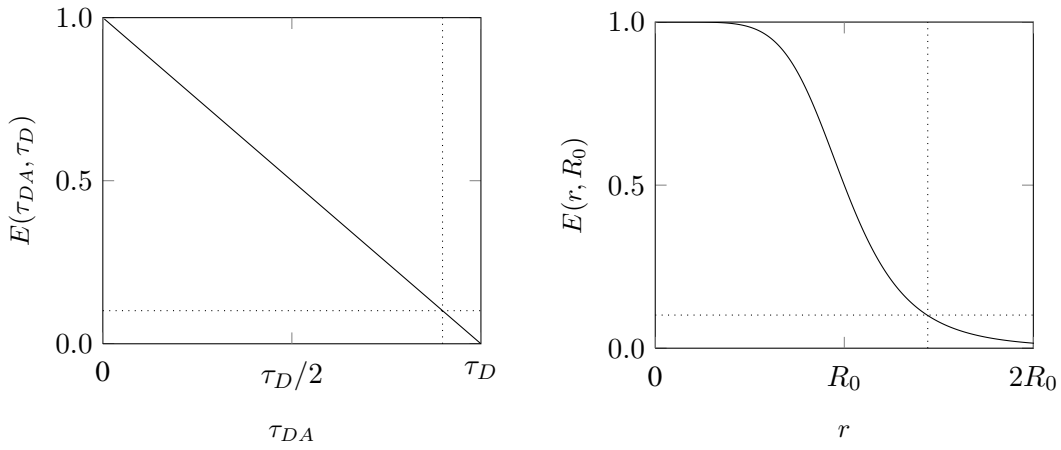
$$E = R_0^6 / \left(R_0^6 + r^6\right), \tag{1.1}$$

where R_0 is the Förster critical distance (the distance at which the FRET efficiency is 50%). A reduction in the donor fluorescence lifetime occurs as a consequence of the FRET process, with the FRET efficiency also being related to the fluorescence lifetime of the donor in the presence of the acceptor τ_{DA} and the donor lifetime in the absence of the acceptor τ_D, by the following,

$$E = 1 - \tau_{DA}/\tau_D \tag{1.2}$$

thereby enabling the separation of the donor and the acceptor to be inferred if the fluorescent lifetimes τ_{DA} and τ_D are known. Careful analysis of FLIM data for the reliable and accurate estimation of the fluorescent lifetimes τ_{DA} and τ_D (and the other parameters describing a fluorescence decay) is crucial to conducting FRET experiments via fluorescence lifetime. The variation of the FRET efficiency with fluorophore separation (Eqn. (1.1)) and with the lifetime of the quenched donor (Eqn. (1.2)) is shown in Fig. 1.1.

Typically, to quantify protein-protein interactions and study biological molecules that are of interest using FLIM, fluorescent probes are introduced into the specimen being



(a) FRET efficiency, $E(\tau_{DA}, \tau_D)$

(b) FRET efficiency, $E(r, R_0)$

Figure 1.1: *FLIM and FRET efficiency for measuring fluorophore separation:* In (a) the FRET efficiency as a function of the fluorescence lifetime of the donor in the presence of the acceptor τ_{DA} and the donor lifetime in the absence of the acceptor τ_D (Eqn. (1.1)), and in (b) the FRET efficiency as a function of the donor fluorophore and acceptor separation r and the Förster critical distance R_0 (Eqn. (1.1)). The dotted lines illustrate that by measuring a FRET efficiency of about 10% through the estimation of the lifetimes τ_{DA} and τ_D it can be inferred that the donor and acceptor are separated by about $1.5R_0$.

studied and their spatial distributions are mapped using FLIM.

The interested reader is directed to the FRET reviews [49, 20, 32, 50, 51], which include discussion of the determination of the Förster distance, FRET probes, methods for applying FRET experimentally, and summaries of some of the various applications of FRET in biomedical research, including the use of FRET to monitor protease activity in (human immunodeficiency virus) HIV, and in the study of DNA structure.

1.2 Acquisition methods: time-resolved, frequency-resolved

The two common techniques for FLIM employ data acquisition in the time-domain (e.g. [52, 53, 54]), using ultra-short repetitive excitation pulses, and in the frequency-domain (e.g. [15, 47, 48]) by means of a sinusoidally modulated excitation light source for il-

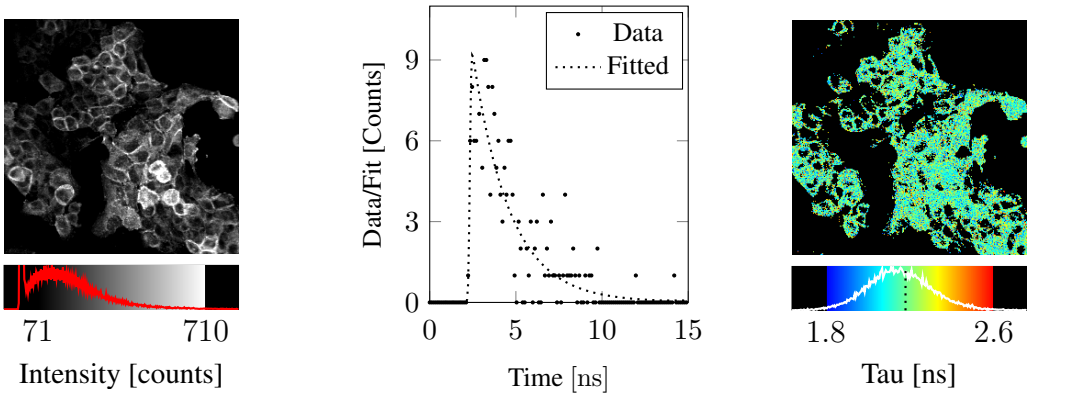
lumination of the sample. Common to both time-domain and frequency-domain FLIM, of course, are excitation sources and detectors; it is how this equipment is applied and the form of the collected data that differ. The general principles behind the collection of time-domain and frequency-domain data, how it is represented and why this influences their analysis, are described here. The different technical solutions used to realise FLIM are reviewed and compared in [44, 45].

1.2.1 Time-domain FLIM

In time-domain FLIM, one of the methods that can be used for data acquisition, and which permits the accurate measurement of the detection time of single photons following periodic pulsed excitation, is time-correlated single photon counting (TCSPC) [55]. It is also possible to perform time-domain FLIM using other methods, such as with a time-gated system e.g. [56, 57] or a using streak camera e.g. [58, 59, 43]. In TCSPC, over time, after many periods of excitation, a data set comprising a collection of photon detection times accumulates that reflects, in this case, a fluorescence decay process that is occurring in the sample being studied. At its simplest, classical TCSPC is achieved by the suitable arrangement of a photomultiplier (PMT), fast discriminators, a time-to-amplitude converter (TAC), an analogue-to-digital converter (ADC), and a memory where the address of each storage location corresponds to a detected time, in order that the detection time of a single photon pulses can be obtained with high precision. A pulsed laser light source provides periodic excitory illumination of the sample causing subsequent fluorescence photon emission; the detection time (relative to the time of excitation) of a single photon pulse is then determined and recorded in memory. The cost, however, of detecting photon arrival times with such high resolution is that the TCSPC hardware necessitates that the likelihood of more than one photon being detected in a single excitation period is negligible so as to avoid pulse pile-up affects [55]. Therefore, over the course of acquiring TCSPC data there are far fewer excitation periods in which a photon is counted than not. Additionally, following the detection of a photon the TCSPC hardware must recover; during this 'dead time' the system is blind and photons cannot be counted [55]. Only one photon event is counted in a single period in TCSPC, and therefore if the likelihood of photon detection is too high early photons will be over-represented in the recorded data; this effect of pulse pile-up is a principal cause of signal distortion [60]. A more detailed

account of the technical details of TCSPC instrumentation and its configuration can be found in [55, 46, 60].

In TCSPC FLIM the so called 'photon distribution mode' of data recording is most commonly employed; the detected photon arrival times are recorded in a histogram (i.e. the order of recording the detected photon arrival times is not preserved), each detection event incrementing the photon count of a particular channel or bin, that is, an interval contained in the overall measurement interval. In photon distribution mode, the accumulation of the arrival times of many detected photons yields a distribution of photon probability that is representative of the fluorescence decay, much like the example shown in Figure 1.2. Alternatively, in the 'time-tag' or 'FIFO' mode, arrival time data for each detected photon is recorded sequentially (i.e. the ordering of detection of counted photons and the experimental time are preserved), again each arrival time being recorded with high temporal resolution. In the context of this work, perhaps the greatest potential advantage of FIFO data recording is that the FIFO storage buffer can be continuously read, coupled with online data analysis this could provide a feedback loop that may permit data acquisition to be terminated once sufficiently confident that enough data has been collected to enable sufficiently accurate parameter estimates.



(a) Intensity image (b) Time-resolved data (c) Lifetime image

Figure 1.2: *Time-domain FLIM:* In (a) an intensity image, in (b) the time resolved photon counting histogram data from a typical pixel, and in (c) the corresponding lifetime image. On application of an appropriate analysis algorithm (at each image pixel) to the FLIM data can be fitted the characteristic fluorescence decay lifetime(s) can be estimated and a fluorescence decay fitted to the recorded data.

A typical modern time-domain FLIM system that excites a sample at repetition fre-

quencies between about 1 MHz and 100 MHz and incorporates a commercially available state of the art TCSPC device, such as the SPC-830[1], can record photon arrival times within an accuracy of about 7 ps (FWHM) and supports a useful count rate of about 4 MHz [60]. Assuming that the TCSPC device is operating at the maximum useful count rate, in acquiring data for, say, a 256×256 image a maximum of only about 60 photons can be counted at each image pixel per second. It is quickly apparent that, using TCSPC FLIM, typically imaging must be performed for between a few seconds and even a few minutes in order to collect sufficient photon counts at each pixel to permit a reliable analysis. Obviously, any efforts to reduce the acquisition time by more effective analysis when photon counts are not plentiful would enable dynamic processes to observed with a greater temporal resolution and is also desirable for faster high-content high-throughput screening applications.

Time-gated data acquisition is another time-domain FLIM technique where the sample undergoes periodic excitation and the detected fluorescence signal is recorded with reference to the excitation. Time-gated techniques are usually targeted at providing fast FLIM data acquisition, and are realised by detector(s) operating at very high count rates to measure the fluorescence signal that occurs within typically between two and eight (e.g. [56, 57, 61]) time-gates that are 'opened' in sequence following excitation. Of course, the practical reality of counting at very high rates for time-gated data acquisition comes with a compromise somewhere. Although, for example, the suitable arrangement of a gated-intensifier phosphor screen and a CCD camera can provide high-count rate detection that evades the main issues that cause TCSPC data acquisition to be slow (namely pulse pile-up and a significant dead-time), in practice, they are not very sensitive and therefore require a strong signal for operation. In [61], using two time-gates and under the assumption of a mono-exponential fluorescence decay at each of the 336×256 image pixels that span the $32 \times 22\ \mathrm{mm}$ imaged field, a wide-field endoscopic FLIM system capable of acquiring FLIM images at up to 29 Hz was demonstrated. In [56], it was demonstrated that both mono- and bi-exponential decays could be accurately measured using four- and eight-gate time-gating systems with non-uniform gate widths, and that, on imaging a sample containing several different fluorescent dyes, lifetime differences of 0.3 ns could be detected.

[1] SPC-830, Becker & Hickl GmbH, Germany

The ‘Galileo’ time-domain FLIM microscope

The in-house developed time-domain FLIM system, currently used at the Randall Division of Cell & Molecular Biophysics, King’s College London, with excitation being provided by a supercontinuum 40 MHz source (SC450-4, Fianium, UK) and photon counting performed using a SPC-830 TCSPC board (Becker & Hickl, Berlin, Germany), is pictured in Fig. 1.3. Considerable details of the design, the various components, development, and some of the research that has been carried out using ‘Galileo’ and other similar in-house developed systems are presented in [62] and on our group website (http://users.ox.ac.uk/∼atdgroup).

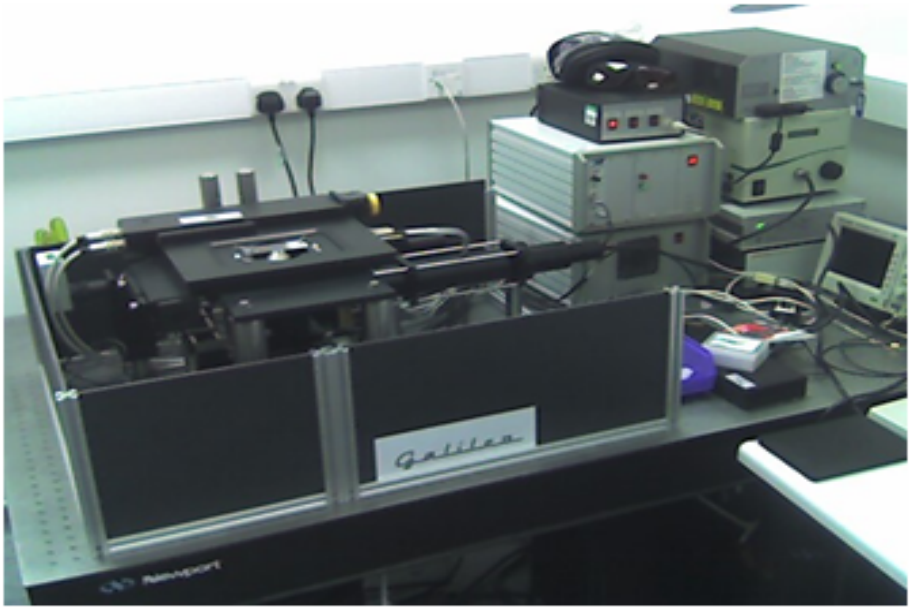


Figure 1.3: *The ‘Galileo’ time-domain FLIM microscope at King’s College London:* Pictured is an in-house developed time-domain FLIM system at the Randall Division of Cell & Molecular Biophysics, King’s College London. Details of the design, the various components, development, and some of the research that has been carried out using ‘Galileo’ and other similar in-house developed systems are presented in [62] and on our group website (http://users.ox.ac.uk/∼ atdgroup).

A schematic of the Galileo fluorescence microscope is shown in Fig. 1.4. A supercontinuum photonic fibre laser produces ‘white light’ (460-2000 nm) pulses (4 ps duration) at a repetition rate of 40 MHz, delivering this through a fibre terminated with a collimating lens; near infra-red components are removed by a pair of ‘hot’ mirrors,

dumping the unwanted energy to a heatsink and using it to provide a reference signal to the acquisition electronics. The resulting laser beam is passed through a beam expander and a pinhole spatial filter before the required excitation wavelength band is selected. The collimated excitation light is reflected by a dichromatic filter onto a galvanometer-driven mirror scanner which scans the beam in the horizontal (line) direction. A pair of relay lenses, acting as an afocal 1:1 telescope, make the vertical (frame) galvanometer scanner axis appear to be at the same point in space as the line scanner. Together these two scanners are focused onto an image plane by the scan lens. The tube lens provides a collimated excitation beam onto the back focal plane of the objective, the horizontal and vertical scans 'pivoting' about the same point in the back focal plane: an objective-focused spot is thus scanned across the sample in the horizontal and vertical directions. The resulting fluorescence is collected by the objective and produces a collimated beam which follows the same path as the excitation beam until the dichromatic reflector, which separates it and guides the fluorescence through a bandpass emission filter. A confocal lens focuses the fluorescence through a confocal aperture and a photomultiplier collects light passing through it. The photomultiplier provides a start signal to the TCSPC acquisition electronics which are arranged to provide imaging information from frame, line and pixel clocks provided by the scanner electronics. The latter can be programmed to provide the desired number of lines and pixels and the require scanning speed, which in turn determines the pixel dwell time. For simplicity, the galvanometer scanners are depicted to be in the same plane; in reality the two are rotated at 90 degrees with respect to each other.

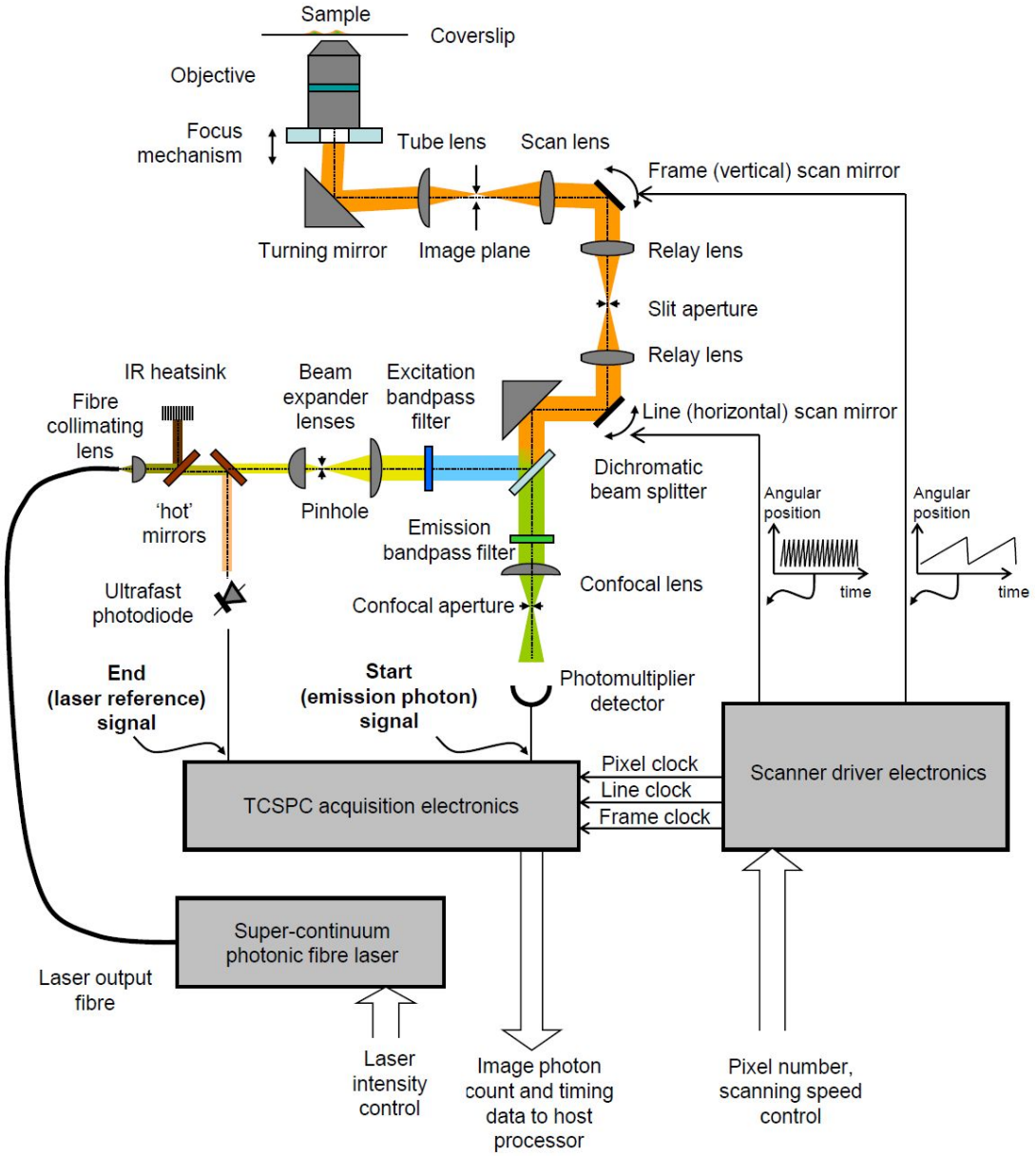


Figure 1.4: *The FLIM acquisition system used on Galileo:* A schematic of the arrangement of the components of the Galileo fluorescence microscope, an in-house developed system [62] currently in use in the Randall Division, King's College London.

1.2.2 Frequency-domain FLIM

In frequency-domain FLIM the sample is excited continuously by a modulated light source to result in delayed fluorescence emission, as illustrated in Fig. 1.5, and fluorescence lifetime(s) can be determined from from the phase and modulation changes [15].

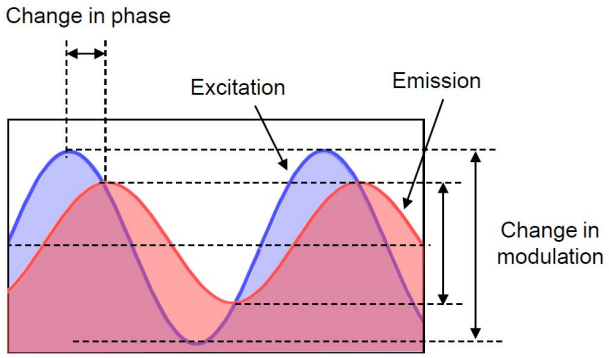


Figure 1.5: *Lifetime determination using frequency-domain FLIM:* By using sinewave-modulated excitation light (blue), of high enough modulation frequency to result in delayed fluorescence emission (red), the fluorescence lifetime can be determined from phase and modulation changes.

The fluorescence lifetime(s) are determined from the demodulation depth, or the reduction in amplitude relative to the mean signal, and the phase lag due to the fluorescence waveform being delayed in time relative to the excitation waveform. The fluorescence lifetime τ of a mono-exponential decay, for example, is related to the demodulation depth M and the phase lag ϕ by the relations,

$$M = \frac{1}{\sqrt{1+\omega^2\tau^2}}, \qquad \phi = \tan^{-1}(\omega\tau) \tag{1.3}$$

where ω is the radial frequency of modulation [63]. The modulated excitation waveform can be of any repetitive form and need not necessarily be sinusoidal as in the data analysis the repetitive signal can be decomposed into a Fourier series and each sinusoidal harmonic component can be treated separately [15], although higher frequency harmonics are effectively removed as the fluorophore acts as an effective low-pass filter [63]. The fluorescence emission signal oscillates at the same high-frequencies as the corresponding

Fourier components of the excitation light, but exhibits a demodulation and a phase lag at each frequency. In acquiring frequency-domain FLIM data, frequency mixing is then employed in order that a gain modulated detector converts the high frequency fluorescence signal (MHz) to a low frequency (10-100 Hz) or constant DC signal which is more easily measured. There are two different frequency-domain FLIM acquisition methods resulting from different approaches to performing frequency mixing, the homodyne method in which the mixing occurs at the exact frequency of the excitation modulation and the heterodyne method in which mixing occurs at a frequency different to, but very close to, the frequency of excitation [63]. In practice, these different detection methods offer two different frequency-domain FLIM imaging modalities.

In frequency-domain FLIM, heterodyne detection is usually used in the scanning mode of imaging, a focused excitation beam being scanned across the sample to collect the image pixels sequentially. The measured signal is a difference, or cross correlation, frequency oscillating at the difference between the light modulation and the detector gain modulation [63]. Typically, in heterodyne frequency-domain FLIM only a single frequency is used to modulate the excitation light at any one time. Wide-field imaging can be realised in frequency-domain FLIM by employing homodyne detection as a gain-modulated image intensifier and CCD detector can be used for where the data for each image pixel is acquired simultaneously [63]. In practice, the detector gain is modulated at the same frequency as the excitation and images are taken at a series of relative phase offsets with the the phase shift and demodulation of the fluorescence emission being determined by fitting a sinusoid to measured data signal.

Chapter 2

Lifetime analysis and fitting

In this chapter FLIM data analysis is considered, from a theoretical perspective in Section 2.1 with a review of the different lifetime analysis methods that have been developed for the time- and the frequency-domain, and in Section 2.2 with respect to the many practical issues which may ultimately determine the success and uptake of an analysis method.

2.1 Signal fitting methods

There exists an abundance of literature pertaining to the acquisition and analysis of FLIM data in both the frequency-domain and the time-domain. In 1991, regarding time-domain fluorescence decay analysis, it was remarked in [64] that 'the multiplicity of the methodologies which have already been developed, suggest clearly the uncertainty investigators still feel regarding the validity of any one method of analysis, irrespective of the statistical criteria used to buttress the results obtained'. A little over two decades have since passed and rather than a single method for FLIM data analysis having become dominant yet more analysis techniques have been introduced and applied to FLIM data, a trend that reflects the increasing popularity and widespread use of FLIM in the biomedical sciences. Despite some recent trends towards global analysis (e.g. [65, 66, 67]), the development of alternative analysis techniques that do not rely on lifetime-fitting (e.g. [68, 69, 70, 71]), and the deployment of alternative models (e.g. [72, 73]) to describe the imaged fluorescence decay processes, a number of competing analysis techniques remain in use, each having their own strengths and weaknesses, and no doubt advocates. It is the

(usually quantitative) description of a fluorescence process that is ultimately of interest, it being unimportant whether the FLIM data has been acquired in the time-domain or the frequency-domain and how they have subsequently been analysed. In this chapter the competing FLIM data analysis algorithms are compared, their robustness and accuracy, ease of implementation, and ease of interpretation all being considered.

In the time-domain, fluorescence decay data acquisition involves the collection of photon emission times that have been counted as occurring within a particular channel (or histogram bin) to yield a histogram of photon count against arrival time; the more photons that are counted the more accurately the histogram represents the fluorescence decay. The most commonly applied analysis methods for time-domain FLIM data involve the direct fitting of a fluorescence decay model to the measured histogram. In this "direct" fitting approach to time-domain FLIM data analysis a fluorescence decay model is fitted directly to the photon counting histogram, the optimal fit being determined according to a goodness-of-fit parameter which somehow quantifies the distance between the fit and the measured photon counting histogram. The key ingredients of the direct fitting approach are usually a measured instrument response function that characterizes the measurement system and a fluorescence decay model which is typically of the form,

$$I(t) = Z + \sum_{\ell=1}^{L} A_\ell e^{-t/\tau_\ell} \tag{2.1}$$

where $I(t)$ represents the fluorescence intensity (the photon count) at time t, Z represents the constant background level, and each of the $\ell = 1, \ldots, L$ decay component is described by an initial intensity A_ℓ and a decay lifetime τ_ℓ. In the direct fitting approach, the optimal fit is determined by finding those parameter values that optimise the goodness-of-fit parameter; in performing such an analysis, iterative reconvolution with a (usually) measured instrument response is typically performed numerically. In applying the direct fitting approach, its statistical effectiveness will depend upon how the goodness-of-fit parameter is defined and its appropriateness to the nature of the time-domain FLIM data being analysed. Much literature is available, as reviewed herein, detailing the intricacies of the the direct fitting approach employing maximum likelihood estimation (ML) and the least squares (LS) method for time-domain FLIM data analysis, and their respective

benefits and drawbacks; a summary of the most pertinent findings follows.

As its name suggests, in the least-squares (LS) fitting approach, the goodness-of-fit statistic is based upon minimising the squared distance between a set of measurements and a proposed fit to those measurements. The LS method of parameter estimation was considered, in 1993, to be the "gold standard" in fluorescence decay analysis [64] and remains widely used today. In essence, all LS methods operate by seeking the minimum in the χ^2 statistic, of which there exist a few different approaches, the standard variants being Pearson's "exact" and Neymen's "reduced". The stability of convergence of Pearson's "exact" χ^2, which minimizes $\sum_i\{(o_i - e_i)^2/e_i\}$, where o_i and e_i are the measured and theoretically expected values of the $i\text{th}$ channel respectively, has been shown to be acutely sensitive to the initial guesses of parameter values [74]. Neyman's "reduced" χ^2 approach minimizes the quantity $\sum_i\{(o_i - e_i)^2/o_i\}$ and appears to be the most commonly used LS approach in fluorescence decay analysis at the moment. However, the effectiveness of the LS approach with small sample sizes has been often questioned due to the assumption of Gaussian statistics in treating noise in the data (e.g. [75, 64, 74]). Ideally, the quantity χ^2 should be weighted according to the standard deviation in each of the measured o_i, that is $\sum_i\{(o_i - e_i)^2/\sigma_i\}$, rather than using the Pearson or Neyman weighting [76], though this is, of course, undefined for empty bins (i.e. those channels where $o_i = 0$), a problem that also exists with Neymen's "reduced" χ^2 statistic[1] and is amplified as the "range of observation or the number of bins increases" [77]. It is claimed in [76], that the issues affecting the accuracy of standard LS predictions for small photon counts "are not due to the Poisson distribution per se but are entirely due to the weighting used" and that regardless of the nature of the noise distribution a systematic bias is present [76]. A means of handling bins that are either empty or have a very low photon count has been investigated in [74], whereby "totalling the count in successive bins improves the estimation in the LS method". The solution proposed in [76] to counteract this systematic bias is to weight the χ^2 statistic using the outcomes of an initial preliminary fit to the data, and it is claimed that such an approach offers a "precision similar to that of ML methods" [76]. A more complicated approach is suggested in [77], having assessed different binning schemes it is stated that "the best we have so far found involves first ordering the times in the data set and then constructing the bin "walls" so as to include a specified number of counts

[1]The usual approach to handling empty bins and bins having a count less than some predefined minimum is to set their count to some predefined minimum.

in each bin". It is concluded in [77] that "with appropriate binning of the data, LS can produce results comparable to those obtained from ML", and, that "the LS scheme may find greater utility in situations where background or multiple decays are present" despite LS being more computationally intensive than ML [77]. It is clear that LS methods, especially for small sample sizes, depend strongly on the channel width. Data manipulation, in order to obtain useful estimates, introduces additional computation and potentially loss of information.

The ML direct fitting approach to the time-domain FLIM histogram is based on the assumption that photon counting obeys Poisson statistics resulting in a goodness-of-fit parameter given by $\chi^2 = 2\sum_i [e_i \log(e_i/o_i) - (e_i - o_i)]$. The ML method for a mono-exponential fluorescence decay was examined in [75] and its performance compared against LS and the "method of moments" estimator. In the case of mono-exponential decay, it was concluded in [75], that "there appears to be no good reason to use a method other than the simple algorithm derived for ML". A very thorough analysis of ML methods for multi-exponential decays was presented a decade later in [64], and tools for the determination of the number of exponential components and how accurately each can be detected were introduced. It is claimed in [64] that ML "provides generally more accurate estimates of lifetimes and fractions than does the standard LS approach especially when the lifetime ratios between individual components are small". More recently, in [74], lifetime and error estimates obtained using the ML method were found to be "almost independent of the total [photon] count" and not to depend on the bin width. However, a hybrid LS-ML analysis is suggested in [74] due to instability in ML algorithms as "convergence of the fitting is critically dependent on the selection of initial guesses of the parameters". A weakness of the ML method is highlighted in [78], which states that "the likelihood does not use information other than the data itself to infer the values of the parameters" and so does not prohibit nonsense estimates as it does not account for such cases whereby "a physical understanding of the circumstances surrounding an experiment can suggest that some values of the parameters are impossible". The ML method and a Bayesian method, which does enable the easy incorporation of prior knowledge, are compared for a mono-exponential decay in [79] using "an exact probability model of the photon arrival times observed with background noise", and although both approaches yield good estimates it was concluded that "while either the Bayesian or ML procedures work well for analyzing fluorescence emissions, the Bayesian methods provide more realistic measures

of uncertainty in the fluorescent lifetimes for any combination of background noise and fluorescence emission".

Of course, regardless of whether using a direct fitting approach based on ML or LS, or any other analysis technique for that matter, if Eqn. (2.1) does not faithfully represent the observed data then quantification of the underlying fluorescence process is unlikely to be successful; obviously, for example, the analysis of bi-exponential decay data with a mono-exponential decay model may not offer any reliable insight into the underlying fluorescence process. The responsibility of determination of the decay order is usually left to the user and is performed by visual inspection of the residuals. An additional problem, as discussed above, is the desire to quantify the parameters of a bi-exponential decay for a FRET-based experiment combined with a lack of photon counts in the available data, a frequently encountered situation in FLIM and one that could lead a user, particularly a non-expert in analysis, to be too optimistic regarding the reliability of estimates that can be extracted from low count data. In applying an overly complicated model to limited data, the risk of over-fitting, that is fitting to noise in the data, is a concern.

Although the user may reasonably expect that the physics of the experimental system have been rigourously modelled and that an analysis should be valid in all circumstances, this is not actually the case for most FLIM analyses presented to date. It is striking that although repetitive excitation in time-domain FLIM is acknowledged in the introductory descriptions of almost all of the referenced papers that discuss time-domain FLIM analysis methods, it is only incorporated formally into the analysis presented in [67] and [80]; its absence from the model being concealed in the assumption that all fluorescence had completely decayed prior to subsequent excitation in the other works. It was noted in [74] that although repetitive excitation can be accounted for, it makes parameter estimation much more difficult. In [72] and [73] a slightly different approach to describing the fluorescence decay process has been adopted, it being proposed that a distribution of characteristic decay lifetimes is more effective for the study of heterogeneous fluorescent populations than a discrete sum of decays each having their own distinct lifetime the discrete sum of decay contributions (as is encoded in Eqn. (2.1)). The stretched exponential approach of [72] incorporates a model of the form,

$$I(t) = I_0 e^{(-t/\tau_{\mathrm{kww}})^{1/h}}, \tag{2.2}$$

to describe a fluorescence decay process, where τ_{kww} is the characteristic timescale[2] of the decay and h is the heterogeneity parameter. It is claimed that the model captures captures the scenario that the decay rate itself is time-dependent and therefore stretched due to the presence of progressively depleted random sinks that capture excitations, but it is also noted that it is important to realise that there are situations that directly lead to a stretched-exponential decay even in the absence of energy transfer mechanisms (i.e. in the absence of heterogeneity).

In applying a direct fitting approach to the measured time-domain histogram data the choosing (or defining) of a goodness-of-fit parameter is equivalent to the imposing of a 'noise model' which describes the statistics of counting photons into bins on the analysis. In applying a LS-based analysis, the noise model is implicitly chosen to be Gaussian, a choice which certainly does not correspond to counting discrete events and only approximates the experimental reality when the number of photons counted is large (where Gaussian statistics approximate Poisson statistics). In applying the ML approach to direct fitting of time-domain data the goodness-of-fit parameter follows from the expectation that the counting of photons into bins obeys Poisson statistics. Clearly, therefore, ML-based direct fitting to the measured histogram is preferable to LS-based fitting, especially at low total photon counts. However, it should be noted that although the application of Poisson statistics to counted photon event data is certainly more appropriate than using Gaussian statistics, it has recently been shown that even Poissonian statistics cannot necessarily be assumed of a counter that has a considerable dead time [81], such as may be the case for some TCSPC hardware. Unfortunately, despite the known limitations of the direct fitting LS-based analysis its use remains widespread in time-domain FLIM, largely down to its ease of implementation and to the fact that it is packaged with popular commercial FLIM analysis implementations. The realisation of a ML-based analysis by the simple adaption of a standard LS-based analysis using the Levenberg-Marquardt algorithm is presented in [82].

In the rapid lifetime determination (RLD) technique (e.g. [57, 61]), the ratios of the fluorescence signal recorded in different time-gates are used to estimate the fluorescence decay parameter values. In the simplest possible case of a time-gated system having two

[2] "Also known as the Kohlrausch-Williams-Watts function, the stretched exponential function was first studied by Kohlrausch in 1847 as an empirical description for the structural relaxation of glassy fibers and subsequently used by G. Williams and D. C. Watts to describe dielectric relaxation in polymers" [72].

gates of equal width and and an interval of ΔT between their respective opening, the lifetime τ of a mono-exponential fluorescence decay is given by,

$$\tau = \frac{\Delta T}{\ln(I_2/I_1)} \tag{2.3}$$

where I_1 and I_2 are the measured (integrated) fluorescence intensities of the early and the late time-gate respectively; notice here that the effects of any background have not been considered. Although very simple in both theory and implementation, the application of such RLD techniques to obtain fluorescence decay parameter estimates very quickly comes at the expense of accuracy; for mono-exponential decay data the simplicity of such analysis being accompanied by an increase of about 30%-40% in the uncertainty of the estimated decay lifetime compared than that obtained using a least squares analysis [83]. In [57] different gating schemes and the number of gates required to achieve a given (expected) uncertainty in fluorescence decay lifetime estimated are discussed, it being found that four- and eight-gate configurations having non-uniform gate widths greatly improve the sensitivity of the analysis and required about half as many total photon counts to obtain an accuracy of 5% than does a two-gate system. Gating schemes for bi-exponential decay parameter estimation have been explored in e.g. [84, 56].

In comparing TCSPC and time-gated data acquisition and conceptually extrapolating from narrow TCSPC bins to wide time-gates, it is certainly worthwhile considering when and whether acquiring data at a very high temporal resolution is advantageous; for example, as memory is finite the subdivision of the TCSPC measurement interval to count photons with a very high temporal resolution comes at the expense of the spatial resolution of the acquired image[3].

In the frequency domain, the details of data analysis depend on the number of frequencies at which data has been collected. In [85] the analysis of the frequency-dependent modulation and phase data collected at multiple frequencies is considered for both a mono-exponential fluorescence decay and a multi-exponential decay; LS analysis being applied to fit sinusoids to the experimental data with the uncertainty in the measured

[3]The SPC-830 TCSPC card (Becker & Hickl GmbH, Germany), operating with 1 detector channel, can record a 256×256 pixel image when the number of time channels (bins) that subdivide the measurement interval is 256, but could offer a 1024×1024 pixel image recording if photons were counted into only 16 time channels instead [60].

phase and modulation data being estimated from the difference between the measured values and those calculated in known conditions.

Global analysis algorithms were utilized initially in frequency-domain FLIM data analysis (e.g. [65, 86, 66]) and have since been developed for time-domain FLIM (e.g. [67, 87, 88]) with the aim of enabling reliable fluorescence decay parameter estimates to be obtained even in poor signal-to-noise conditions by exploiting the expected spatial invariance of some of the properties of fluorescence decays across an image. The global analysis algorithms presented in [65] and [67], for application in the frequency-domain and the time-domain respectively, are both based on the assumption that the number of fluorescent molecule species present in an imaged sample is limited, and that their lifetimes are not expected to vary spatially. The incorporation of such an assumption into the analysis of poor signal-to-noise data should provide a means of counteracting modest photon counts and significant noise in the collected data at each individual pixel. In frequency-domain FLIM, an additional and significant advantage over conventional frequency-domain analysis techniques of using a global analysis method is that the two lifetimes of a bi-exponential fluorescence decay can be determined even when using a single-frequency FLIM setup; thereby enabling FRET to be studied quantitatively from single-frequency frequency-domain FLIM data [65]. In [65] a global analysis approach was described that simultaneously fits the fractional contributions in all image pixels and the spatially invariant lifetimes. A global analysis algorithm for time-domain FLIM is presented in [87], "that is fast, provides quantitative results, ...and robustly handles time-resolved images with low photon counts" [87], one particular advantage of such an analysis method is noted; "the fact that interacting fractions can be derived from transients of only several hundred counts compared with several tens of thousands is indicative of the power of the technique" [87].

In [89], a lifetime moments analysis (LiMA) is developed for frequency-domain FLIM, providing a quantitative measure of lifetime heterogeneity. An estimator of the width of the lifetime distribution is developed, offering insight into the lifetime distribution heterogeneity in analogous way to the stretched exponential approach of [72] for time-domain FLIM (Eqn. (2.2)). The analysis also demonstrates analytically that the phase-determined and modulation-determined lifetimes (Eqn. (1.3)) diverge for a heterogenous lifetime distribution, a commonly reported experimental observation that had previously had neither a "clear theoretical explanation or analytical solution".

Phasor analysis is a technique that has its origins in the analysis of frequency-domain FLIM data (e.g. [90, 69]) and has since also been applied to time-domain FLIM data in [70, 91]. The primary motivation behind the introduction of phasor analysis to FLIM is that it is considered to offer a more intuitive representation of the experimental process and this may have particular appeal with the non-expert user. In [70] the phasor approach was applied to time-domain FLIM data, the photon counting histogram data being translated to the coordinate $(g(\omega), s(\omega))$ in phasor space by means of the following transformations,

$$g(\omega) = \frac{\int_0^\infty dt\ I(t)cos(\omega t)}{\int_0^\infty dt\ I(t)}, \qquad s(\omega) = \frac{\int_0^\infty dt\ I(t)sin(\omega t)}{\int_0^\infty dt\ I(t)} \tag{2.4}$$

where $I(t)$ is the fluorescence decay model and ω is the excitation angular repetition frequency. In performing FLIM data analysis using the phasor approach, the FLIM data for each image pixel are mapped to a point in phasor space to create a phasor plot. A user would then typically inspect the phasor plot for clustering and dispersion of phasors, fluorescence lifetimes being interpreted with reference to the the so-called "universal circle" in phasor space, a half-circle on which the lifetimes of all purely mono-exponential decays of the form $I(t) = Ae^{-t/\tau}$ lie and within which the phasors of multi-component decays reside. Although the graphical presentation of the FLIM image data that phasor offers may well appeal to both the expert and non-expert user, it should also be noted that in applying the phasor transforms (Eqn. (2.4)) to time-domain data using a decay model of the form $I(t) = Ae^{-t/\tau}$, as is the case in [70], nowhere in the analysis is the influence of the instrument response or a possible non-zero background signal acknowledged. This may be only a minor issue if, across the whole FLIM image (or a sequence of images should they be being compared), the instrument response is uniform and narrow and any background count is negligible; something that may be solely true in laboratory-based studies performed using a fixed microscope system but something that could not necessarily be assumed and may more difficult to accomplish in, say, the potential clinical applications of endoscopic FLIM. A supposedly complementary method of quantitative FRET analysis that does not require lifetime fitting designed for use at high spatiotemporal resolution is presented in [71], where the mathematical minimization of the quantity of donor molecules undergoing FRET yields a new quantity, the minimal percentage of

donor molecules involved in FRET that can be estimated using the average fluorescence lifetime, hence "allowing fast acquisitions to be performed" [71]. This method is, according to [71], complementary to phasor analysis. In considering phasor analysis, it is worthwhile recalling once more that fluorescence lifetime is usually used as a proxy to study some other (biological) property of interest and that the user is unlikely to be an expert in the analysis of FLIM data, something that may be especially so if FLIM becomes more widely employed in clinical diagnostics. Despite its noted drawbacks for non-ideal time-domain data phasor analysis has been developed primarily to present the data from FLIM experiments in a way that should appeal to and be easily accessible to a non-expert (in analysis) user. Certainly, the representation of analysis results to a user (and especially a non-expert user) and how easily they can be interpreted should not be overlooked.

Another approach that obviates the requirement for fitting and is based on the Laguerre expansion has been presented in [68, 92]. A FLIM method based on the Laguerre expansion technique was proposed in [68], where it is stated that a "unique characteristic of this approach is that it can reconstruct a fluorescence response of arbitrary form". It is noted in [68], however, that "although the results of the present study are encouraging, ...the proposed Laguerre method still need[s] to be thoroughly validated on a broad variety of FLIM applications", where it is also observed that "the Laguerre method needs only half or less of the acquired delayed images available for the analysis of the measured FLIM data". Enhancements to the Laguerre analysis of [68] are proposed in [92], offering "a fully automated deconvolution method for TRFS [time-resolved fluorescence spectroscopy] data analysis based on an iterative Laguerre expansion approach" and it is believed the "method will facilitate the use of TRFS in applications where online data analysis is required" [92].

The quantification of exponential decay processes is, of course, not peculiar to FLIM. In employing nuclear magnetic resonance (NMR) techniques for studies in chemical and biological sciences it is also usually necessary to quantify multi-exponential decays. In [93] a Bayesian approach to exponential decay parameter estimation for NMR data is presented and in the related work, [94], the analysis is extended to provide a model selection algorithm that can be used to determine quantitatively the number of exponential components that contribute to a decay. Applied to a transverse relaxation study on rat brain the model selection algorithm predicted that the relaxation data be due to a bi-exponential process, in keeping with expectations from the literature for such a ^{23}Na relaxation study

and also from a qualitative analysis of the residuals [94]. In [95], a Bayesian approach was applied for the purpose of burst identification in burst integrated fluorescence lifetime (BIFL) experiments; the developed algorithm being able to quantify the probability that some data be due to a background signal and probability that the data be due to a fluorescence burst, and was demonstrated to successfully identify fluorescence bursts from simulated data having a 90% background (i.e. decay signal of only 10%). Bayesian analysis has been successfully applied to the processing of biomedical data [78], and for model selection and parameter estimation in systems biology (e.g. [96]). In [97], Bayesian techniques have been applied to infer the composition of multiplexed Raman spectra data, such as would be obtained in disease identification by using labelled DNA sequences.

Of course, no matter how well any estimation technique can perform and how accurately it can analyze the properties of a fluorescence decay, in practical situations it is wise to be mindful that there are numerous ways in which the acquired decay data can be contaminated during its acquisition. The most obvious data contaminant arises from the various delays, reflections, and data spreading inherent to, and at, any interfaces between the experimental apparatus; their collective effect being termed the "instrument response function". In the above analyses, the effects of instrument response are handled by deconvolution, either analytically in the model itself by using a function that approximates the properties of the instrument response or more commonly numerically in the implementation of the estimation algorithms using the measured instrument response function. However, far more subtle artefacts introduced by the TCSPC system should also be considered, such as "dead time", which "does not only limit the maximum count rate [...], it can also introduce distortions of the signal waveform and errors in the measured intensities" [60]. In a poorly configured TCSPC system, the dead time can introduce "pile-up" ([98, 99, 60]), such that "if the detection rate is so high that the detection of a second photon within the recorded time interval becomes likely the signal waveform is distorted" [60]. Aside from the optimal configuration of a TCSPC system for a particular experimental purpose, there are also other sensible measures that should be considered to give the employed estimation technique the greatest chance of yielding useful estimates, such as ensuring the measurement interval is well matched to the anticipated fluorescence decay. It is suggested in [74] that (for LS analysis, at least) "the region of the multi-component fitting should be sufficiently larger than the slowest decay time", a finding that is echoed in [100], which recommends that "in the presence of background, it is better to choose a

larger measurement window to increase the significance of the last bins, here of about 30 lifetimes".

Quite a number of different analysis methods have been discussed above, some of the articles cited discuss how well or how poorly the presented analysis performs when there is little data available. A convenient means of quantify, with just one statistic[4], how accurate the estimates that a particular analysis may offer was introduced in [57]; the figure of merit $F = (\Delta\tau/\tau).\sqrt{I}$ providing a measure of the sensitivity of the lifetime analysis method, where $\Delta\tau$ is the standard deviation of the determined lifetime τ, and I is the number of photon counts (in the data yielding the lifetime estimate). In [101], the photon economy and acquisition time are studied for a variety of both frequency- and time-domain systems, and is used in this thesis (Section 6.1.1) for the comparison of the developed Bayesian analysis with other methods.

It is evident, having surveyed much of the literature, that the effective analysis of fluorescence decay is not a simple task. In the analysis of time-domain FLIM data, problems are fewest for large sample sizes, where Gaussian statistics closely approximate Poisson statistics, and both LS and ML have been shown to perform well (e.g. [75, 64, 77, 74]). However, for small sample sizes, standard LS has been shown to yield poor estimates (e.g. [75, 64, 74]), though data manipulation prior to analysis ([77, 74]) and a modification of the LS weighting [76] have both been shown to improve LS estimates to the extent that they are comparable with those from the ML method. It is clear that the analysis of FLIM data is rarely a trivial problem, particularly so in practice where high expectations meet low photon count data. It is also evident that the armoury of FLIM data analysis methods would be strengthened by the introduction of an algorithm that offers a clarity in the assumptions made and a greater opportunity to include physical quirks of acquisition. A method which offers robust error estimates and the joint probability distribution function over the model parameters would be especially useful when data is in short supply.

In Part II a Bayesian method for the analysis of time-domain FLIM data is developed, incorporating a fully analytic time-domain FLIM system model which accounts for the effects of repetitive excitation and includes an analytic approximation to model the influence of the instrument in the data. In developing such a Bayesian analysis for

[4]It should be noted that although the photon economy statistic F can be extremely useful for comparing how efficiently different analysis methods use the available data, in selecting a method it would be wise to confirm that the parameter estimates do not exhibit any systematic bias as such would not be apparent on inspection of F alone.

fluorescence decay parameter estimation the probability distribution of all different combinations of decay parameter values becomes accessible. Not only are the most probable decay parameters values available by locating the peak of the probability distribution but also the probability distribution can be explored and the relative likelihoods of different combinations of fluorescence decay parameter values can be assessed probabilistically. Additionally, a reliable measure of the uncertainty in any of the parameter estimates becomes possible, something of great importance if parameter estimates are to be used with confidence.

2.2 Expectations and practicalities

It is clear that, whether using time-domain or frequency-domain, at some point in a FLIM experiment it becomes necessary to somehow extract the fluorescence lifetime(s) from the collected data. The fluorescence lifetime(s) itself is unlikely to be a quantity of real interest in an experiment, but rather acts as a proxy for some biological or chemical property that is ultimately of interest (and that cannot usually be measured directly as effectively inferred from other methods). The effectiveness of an experiment to quantify, say, protein-protein interactions using FRET therefore depends crucially on the effectiveness of the analysis method employed to extract fluorescence lifetimes from FLIM data. Although this work is focussed mainly on the theoretical and technical development of a Bayesian analysis for FLIM data, it is worthwhile considering from the point of view of an experimentalist where data analysis sits in the experimental chain. In this section, some of the practical and experimental issues that may be helped or hindered by an analysis method and its implementation are highlighted.

Analysis of FLIM data is a statistical problem. Many different methods of analysis have been previously studied and are applied to experimentally measured data to quantify fluorescence decay in FLIM experiments. An ideal analysis method would be able to produce extremely accurate and easily interpreted parameter estimates, be quick and be easy to use, whilst also being easy to implement and conceptually easy to understand. In reality, of course, no analysis or algorithm is likely to provide all of these features and it is necessary for certain requirements to be prioritised at the expense of others.

Of course, it is desirable, if not essential, that an analysis method should always provide reliable fluorescence decay parameter estimates. However, it is understood that the

accuracy of any analysis method and its fluorescence decay parameter estimates is limited by how much data are available for the analysis. The greatest accuracy possible from analysis of the available data is always desirable. In a FLIM experiment, however, sometimes the time taken for an analysis to yield results may be more important than those results being of best possible accuracy [102]. An experiment in which online analysis provides a critical feedback loop, for example, would be such a situation where the accuracy of parameter estimates must be sacrificed in favour of speed of analysis (regardless of the amount of data available); a slow analysis algorithm would not only be irksome to an impatient user but also of no use for online data processing in such experiments. Conversely, it may be that the most accurate results possible are desired when only very limited data is available, such as if attempting to resolve dynamic processes that have a time scale similar to an imaging duration that yields extremely challenging data. Of course, just as the level of accuracy required and the time taken for analysis will differ from experiment to experiment, the most appropriate analysis technique will also depend upon the requirements of the experiment.

It is certainly also desirable that any analysis provide some measure of how much confidence can be placed in the parameter estimates. Aside from the integrity of experimentally obtained data, the measure of confidence in any estimated model parameter values depends also on the particular data modeling approach utilized and the potentially many nuances that are (intentionally or unintentionally) unaccounted for in any theoretical modeling of a physical process. It is imperative that the quality, or goodness, of any fit to measured data can be critically assessed rather than a particular model and its parameter predictions being accepted with blind faith. Should such critical assessment be somehow inherent to the analysis technique and not require user interaction or inspection then all the better. Regarding typical FLIM imaging, an analysis that could not only yield parameter estimates and a measure of their accuracy and reliability, but do so sufficiently quickly as to enable online data processing that indicates when sufficient data has been acquired at an image pixel could reduce imaging durations and also reduce the exposure of the sample being studied to (excitation) radiation.

As considerable effort and time is expended in the experimental process prior to and subsequent to acquiring FLIM data, it is wise to spend time in choosing the right analysis method. Most importantly, it is imperative that the wrong analysis is not applied to some FLIM data and the potentially erroneous results then used in the formation of

equally potentially erroneous conclusions regarding biology or chemistry (or whatever) being studied.

Ultimately, though, a user will decide on which method to employ in the analysis of their FLIM data. Sometimes, unfortunately, it may well be that factors far removed from the accuracy of a methods estimates and its statistical robustness play a large part in leading a user to adopt a particular FLIM analysis technique, particularly in the setting of a busy multi-disciplinary laboratory. The ease of use[5] of an algorithmic implementation[6] of an analysis technique could certainly prove persuasive, as could the ease of interpretation of the resulting estimates. The speed with which an analysis can be performed is also usually of considerable importance; many current studies that employ FLIM amass large quantities of data and sometimes thousands of FLIM images must be analysed. Although perhaps to the theorist these could be seen to be largely irrelevant concerns they are likely to be of significance to the experimentalist and will ultimately have a bearing on whether an algorithm gains traction in the community and enjoys widespread use or not regardless of how good an analysis method may (or may not) be.

[5] Any requirement of user interaction in an implementation, for example, for the technical configuration of an algorithm is more likely to be greeted as a complicating deterrent than a welcome freedom in the analysis.

[6] The availability, cost, and ease of installation are all likely to be important in determining which software package/analysis method a laboratory uses; although these issues aren't of direct concern to this work their importance should be at least noted.

Part II

Probabilistic fitting: theory

The primary purpose of this work is the quantitative characterization of fluorescence decay processes. It is not usually possible to measure directly a set of parameters that adequately and accurately describe a decay. However, by applying appropriate analysis to data that can be directly measured, in this current work a set of photon counting events, inferences regarding the nature of a decay process are possible. In following such a methodology the intention is to be able to provide answers to a question such as *"What can we say about the underlying processes that generated these data?"*. It is also natural to ask questions like, *"What can be based on what has been said?"*, which is of particular relevance when the decay characterization is not the end goal of a study, where the results of the characterization are built upon in further enquiries and subsequent inference. *"Should I have confidence in these predictions?"* and *"Could the analysis of my data be better?"* are perfectly reasonable (if not critical) questions to which Bayesian analysis is able to provide quantitative answers. Bayesian analysis provides a systematic method of relating data that can be directly measured to the underlying process generating the data.

The tools of Bayesian analysis are introduced in Chapter 3, in roughly the same order as the questions above are posed, by firstly examining parameter inference and subsequently discussing model selection and optimization. In introducing Bayesian techniques in this order their hierarchical nature and versatility become apparent. At the core of the Bayesian analysis for time-resolved FLIM data presented in this work is the modelling of the time-domain FLIM system, such as the system described in Section 1.2.1. In Chapter 4, a time-domain FLIM system model relating the arrival time of a photon counting event to a decay process *and* the measurement instrumentation is introduced. The developed FLIM system model differs fundamentally from those used in existing analysis techniques in that periodic excitation is rigorously incorporated from the outset and that it is fully analytic. The Bayesian methodologies introduced in Chapter 3 are applied in Chapter 5 for fluorescence decay parameter inference, and through exploiting the combination of the hierarchical properties of Bayesian analysis and the fully analytic FLIM system model, additionally for the purposes of fluorescence decay model selection and (FLIM system) instrument response determination.

Chapter 3

Bayesian analysis

Bayesian analysis provides a systematic method for relating data that can be directly measured to the underlying process generating the data. At the root of the Bayesian analysis techniques introduced lies conditional probability theory and, as such, the majority of this work is concerned with describing, discussing, and investigating the likelihood of a parameter having a particular value conditioned on some data that, having been measured, is known. The probability of an event A conditioned on (or given that) that the occurrence of event B is known is denoted by $p(A|B)$ and is termed the conditional probability of A given B.

3.1 Parameter inference

Consider a model having parameters $\boldsymbol{w}$ and a data set $D = \{x_1, \ldots, x_p\}$, applying Bayes' theorem the posterior probability distribution of the model parameters $\boldsymbol{w}$ conditioned on the data D can be expressed in terms of the the likelihood of the data given a set of model parameters:

$$p(\boldsymbol{w}|D) = \frac{p(\boldsymbol{w})p(D|\boldsymbol{w})}{\int d\boldsymbol{w}'\, p(\boldsymbol{w}')p(D|\boldsymbol{w}')} = \frac{p(\boldsymbol{w})\prod_{\mu=1}^{p} p(x_\mu|\boldsymbol{w})}{\int d\boldsymbol{w}'\, p(\boldsymbol{w}')\prod_{\mu=1}^{p} p(x_\mu|\boldsymbol{w}')} \tag{3.1}$$

where the prior distribution $p(\boldsymbol{w})$ encodes any prior knowledge regarding the model parameters $\boldsymbol{w}$ and the data likelihood given the model parameters, $p(D|\boldsymbol{w})$, with the data

assumed to be generated independently (i.e. $p(D|\boldsymbol{w}) = \prod_{\mu=1}^{p} p(x_\mu|\boldsymbol{w})$). Defining the quantity $S(\boldsymbol{w}, D) = -\ln[p(\boldsymbol{w})p(D|\boldsymbol{w})]$, the posterior can be written in the alternative form:

$$p(\boldsymbol{w}|D) = \frac{e^{-S(\boldsymbol{w},D)}}{\int d\boldsymbol{w}'\, e^{-S(\boldsymbol{w}',D)}} \tag{3.2}$$

The most likely parameter values $\boldsymbol{w}^\star$ are those for which the quantity $p(\boldsymbol{w}|D)$ is greatest, or equivalently those that give the minimum of $S(\boldsymbol{w}, D)$, such that:

$$S(\boldsymbol{w}^\star, D) = \underset{\boldsymbol{w}}{\operatorname{argmin}}\, S(\boldsymbol{w}, D) = \underset{\boldsymbol{w}}{\operatorname{argmin}}\, \{-\ln[p(\boldsymbol{w})p(D|\boldsymbol{w})]\} \tag{3.3}$$

The posterior distribution (Eqn. 3.1), however, encodes the parameter likelihood across the entire parameter space and, therefore, considerably more than just the most probable parameter values can be accessed. The posterior marginal distribution for any model parameter can be obtained:

$$p(w_i|D) = \frac{\int d\boldsymbol{w}/dw_i\, p(\boldsymbol{w})p(D|\boldsymbol{w})}{\int d\boldsymbol{w}'\, p(\boldsymbol{w}')p(D|\boldsymbol{w}')} \tag{3.4}$$

where $w_i \in \boldsymbol{w}$ and $\int d\boldsymbol{w}/dw_i$ represents the integral over all model parameters $\boldsymbol{w}$ bar the parameter w_i. In much the same manner, parameter average values $\langle w_i \rangle$ and standard deviation Δw_i can be determined to yield predictions of the form $w_i = \langle w_i \rangle \pm \Delta w_i$, where:

$$\langle w_i^\ell \rangle = \frac{\int d\boldsymbol{w}\; w_i^\ell\, p(\boldsymbol{w})p(D|\boldsymbol{w})}{\int d\boldsymbol{w}'\, p(\boldsymbol{w}')p(D|\boldsymbol{w}')}, \qquad \Delta w_i = \sqrt{\langle w_i^2 \rangle - \langle w_i \rangle^2} \tag{3.5}$$

3.2 Model selection

In performing parameter inference as above, the role of the assumed model has not been explicitly stated. The posterior distribution of Eqn. 3.1 informs of the likelihood of some

parameter values $\boldsymbol{w}$ given some data D. However, it could be considered that there is also some hidden conditioning in Eqn. 3.1; the posterior distribution actually informs of the probability of parameter values $\boldsymbol{w}$ conditioned not only on the data D but additionally on the model that has been assumed. In this section, model selection is introduced, the Bayesian formalism being employed to enable the most suitable model to be systematically chosen from an ensemble of candidate models. Denoting by $\mathcal{H}_K$ the model having parameters $\boldsymbol{w}_K$, the implicit assumption (Eqn. 3.1) of using a particular model in parameter inference is here formalised explicitly:

$$p(\boldsymbol{w}_K|D,\mathcal{H}_K) = \frac{p(\boldsymbol{w}_K|\mathcal{H}_K)p(D|\boldsymbol{w}_K,\mathcal{H}_K)}{\int d\boldsymbol{w}'_K\, p(\boldsymbol{w}'_K|\mathcal{H}_K)p(D|\boldsymbol{w}'_K,\mathcal{H}_K)}$$

The Bayesian framework can be utilised to enable an informed selection as to the most suitable model to be employed in any subsequent parameter inference, the quantity $p(\mathcal{H}_K|D)$ informing of the likelihood of model $\mathcal{H}_K$ given the data D:

$$p(\mathcal{H}_K|D) = \frac{p(\mathcal{H}_K)\int d\boldsymbol{w}_K\, p(\boldsymbol{w}_K|\mathcal{H}_K)p(D|\boldsymbol{w}_K,\mathcal{H}_K)}{\sum_{K'} p(\mathcal{H}_{K'})\int d\boldsymbol{w}_{K'}\, p(\boldsymbol{w}_{K'}|\mathcal{H}_{K'})p(D|\boldsymbol{w}_{K'},\mathcal{H}_{K'})} \tag{3.6}$$

with the model itself now being the subject about which inference is made. In arriving at Eqn. 3.6, the relation $p(D|\mathcal{H}_K) = \int d\boldsymbol{w}_K\; p(\boldsymbol{w}_K|\mathcal{H}_K)p(D|\boldsymbol{w}_K,\mathcal{H}_K)$ has been used to remove the model parameters $\boldsymbol{w}_K$ where again the data likelihood given the model $\mathcal{H}_K$ and its parameters $\boldsymbol{w}_K$ assumes that the data has been generated independently (i.e. $p(D|\boldsymbol{w}_K,\mathcal{H}_K) = \prod_{\mu=1}^{p} p(x_\mu|\boldsymbol{w}_K,\mathcal{H}_K)$). It is possible to go yet further in determining the most suitable model to describe some data D. The optimal hyperparameter value(s) can also be sought using Bayesian techniques and this should certainly be considered for small data sets when the influence of the prior distribution in most pronounced. Denoting by $\boldsymbol{\alpha}_K$ the hyperparameters of model $\mathcal{H}_K$, their influence in conditioning the parameter estimates is easily seen:

$$p(\boldsymbol{w}_K|D,\mathcal{H}_K,\boldsymbol{\alpha}_K) = \frac{p(\boldsymbol{w}_K|\mathcal{H}_K,\boldsymbol{\alpha}_K)p(D|\boldsymbol{w}_K,\mathcal{H}_K,\boldsymbol{\alpha}_K)}{\int d\boldsymbol{w}'_K\, p(\boldsymbol{w}'_K|\mathcal{H}_K,\boldsymbol{\alpha}_K)p(D|\boldsymbol{w}'_K,\mathcal{H}_K,\boldsymbol{\alpha}_K)} \tag{3.7}$$

Again, using the Bayesian framework, the most appropriate model to describe some data D can be determined and optimized simultaneously:

$$p(\mathcal{H}_K, \boldsymbol{\alpha}_K|D) = \frac{p(\mathcal{H}_K)p(\boldsymbol{\alpha}_K)\int d\boldsymbol{w}_K\, p(\boldsymbol{w}_K|\boldsymbol{\alpha}_K)p(D|\boldsymbol{w}_K, \mathcal{H}_K)}{\sum_K' p(\mathcal{H}'_K)p(\boldsymbol{\alpha}'_K)\int d\boldsymbol{w}'_K\, p(\boldsymbol{w}'_K|\boldsymbol{\alpha}'_K)p(D|\boldsymbol{w}'_K, \mathcal{H}'_K)} \tag{3.8}$$

In deploying the model selection and optimization techniques of this section, the main problem that is usually encountered is the computation of the integrals. Fortunately, however, using suitable approximations as to the shape of the distribution under the integral, such difficulties are at least eased.

Chapter 4

Modelling a time resolved FLIM system

In this chapter a model that describes the FLIM system shall be developed. In time-domain FLIM, a laser is usually used to periodically excite a sample, causing fluorescence emission. Fluorescence decay photons are subsequently detected by a photon-counting detector; their arrival times (relative to the pulsed excitation) being recorded with high accuracy. Over time, a set of photon arrival times that represent the fluorescence emission, potentially , accumulates; it is such sets of arrival times that form the 'data' in our Bayesian analysis of time-resolved FLIM data that will be introduced in Chapter 5. In parameterizing the time-domain FLIM system for use in the Bayesian analysis of time-resolved FLIM data, a model that captures the characteristics of the FLIM system as accurately as possible, and relates a particular photon arrival time to a set of fluorescence decay and other model parameters (e.g. for characterisation of the system itself, etc.) is developed. The time-domain FLIM system model assumes the role of the quantity $p(x_\mu|\boldsymbol{w})$ of Eqn. (3.1) of the general introduction to Bayesian analysis of Chapter 3, where in this case the x_μ being a particular (binned) photon arrival time and $\boldsymbol{w}$ a set (or more usually a subset) of FLIM system model parameters. In the interest of developing a model that remains generally applicable, those elements of the FLIM system that exert little influence on the behaviour of typical time-domain FLIM system (and are easily controlled experimentally through appropriate configuration of the FLIM system hardware) are neglected; these are discussed in Section 4.5.

Considering the time-resolved FLIM system in the most simple terms, as illustrated in Fig. 4.1, a sample undergoes repetitive excitation (repetition period T_m), emitting pho-

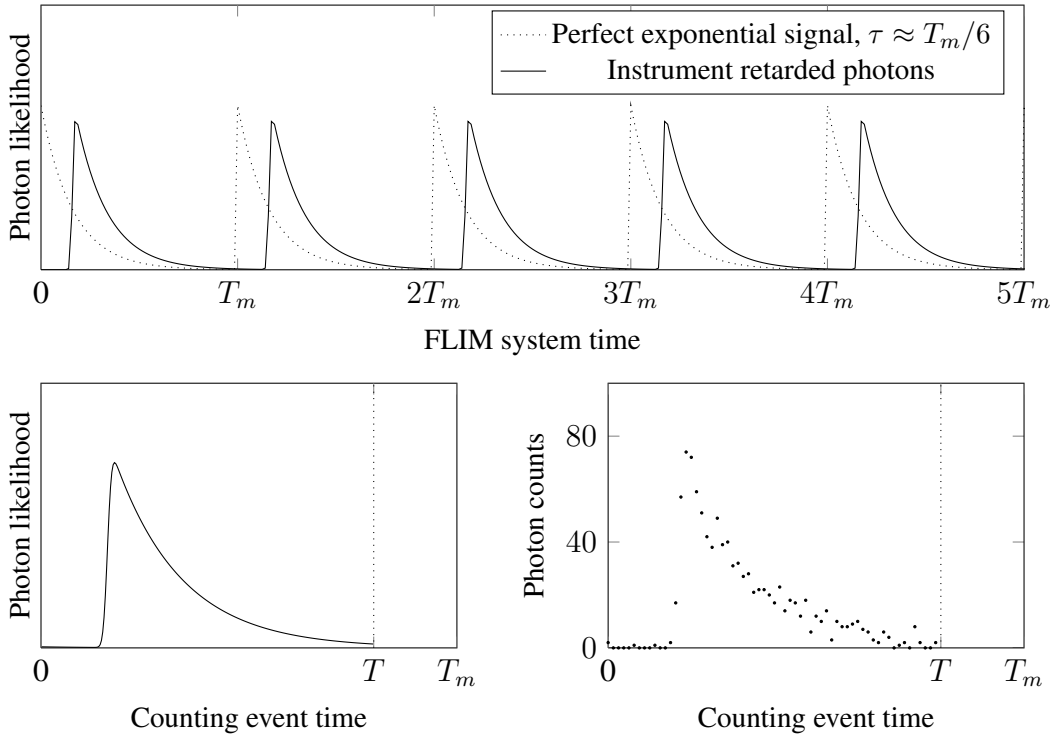


Figure 4.1: *The time-domain FLIM system; repetitive excitation, periodic fluorescence decay, and photon counting:* An illustration of the key components of the time-domain FLIM system model for a FLIM system with repetition period T_m and a measurement window of duration T, a typical instrument response, and a mono-exponential decay of lifetime $\tau \approx T_m/6$. In the upper panel, periodic decay (dotted decays) as a consequence of a sample being subjected to repetitive excitation at every nT_m $(n \geq 0)$ in FLIM system time, with the fluorescence photons being retarded and broadened on progressing through the FLIM system apparatus (solid decays). In the bottom left panel, the fluorescence decay photon detection likelihood within the measurement interval $[0, T]$; notice that there is no likelihood of detecting a photon outside of the measurement interval (i.e. in the interval $(T, T_m]$). In the bottom right panel, typical photon count data for such a decay having about 1000 total photon counts, with photons being counted into 64 bins of equal width subdividing the measurement interval.

tons which traverse a path through some experimental apparatus, before they are detected with a specific quantum efficiency. It is neither desirable nor likely to be advantageous to attempt to model independently the influence of all of the different components of a time-domain FLIM system, such as that shown in Figures 1.3 and 1.4; instead, the effects of the various components of the experimental system are considered here only in that it

is assumed that the instrument introduces a delay of duration u to a signal, distributed according to $\Gamma(u)$. Any arrival time Δt of a detected photon is determined with reference to the most recent excitation pulse[1], and is recorded as having being detected in an interval (i.e. a bin). The following sections add details to the above description, beginning with *what goes in to the system* (i.e. repetitive excitation), followed by *what comes out of the system* (i.e. photon arrival times recorded in discrete time), and *what happens inbetween* (i.e. sample signal responds to the excitation and instrument delays the signal). In the interest of readability, wherever possible only the key assumptions that guide the model development and significant results along the way are presented here; all intermediate steps, technical and mathematical details can be found in full in Appendix A. In this analysis, the events that are analysed are photon arrival times that have been *detected* in the measurement interval $[0, T]$, consequently the likelihood expressions that are normalised over this interval.

4.1 Repetitive excitation

In modelling an experimental system based on periodic excitation and in which any collected data consists of event times recorded on a periodic window, it is crucial that the relationship between the *recorded* data and the *underlying* data generating process(es) be captured. In the FLIM system considered here, it is the arrival time Δt of a detected photon that is recorded on a periodic time window of duration T_m (assumed to be aligned with the periodic excitation), and it is the nature of an underlying fluorescence emission process that is sought (see Fig. 4.2). Formalising this, the recorded photon arrival time Δt determined with respect to the most recent excitation is given by:

$$\Delta t = t + u - T_m.\mathrm{int}\left(\frac{t+u}{T_m}\right) \tag{4.1}$$

where t is the actual *emission* time of a photon due to the fluorescence decay process, the *delay* time u describes the latency between the actual emission process and that recorded by the electronic time measurement system, and (with $\mathrm{int}(z)$ being the largest integer $n \leq$

[1] In practice, using reverse-start-stop TCSPC, it is actually the time between photon detection and the next excitation that is measured, it then being trivial to represent such times with reference to the preceding excitation pulse as the repetition period is known.

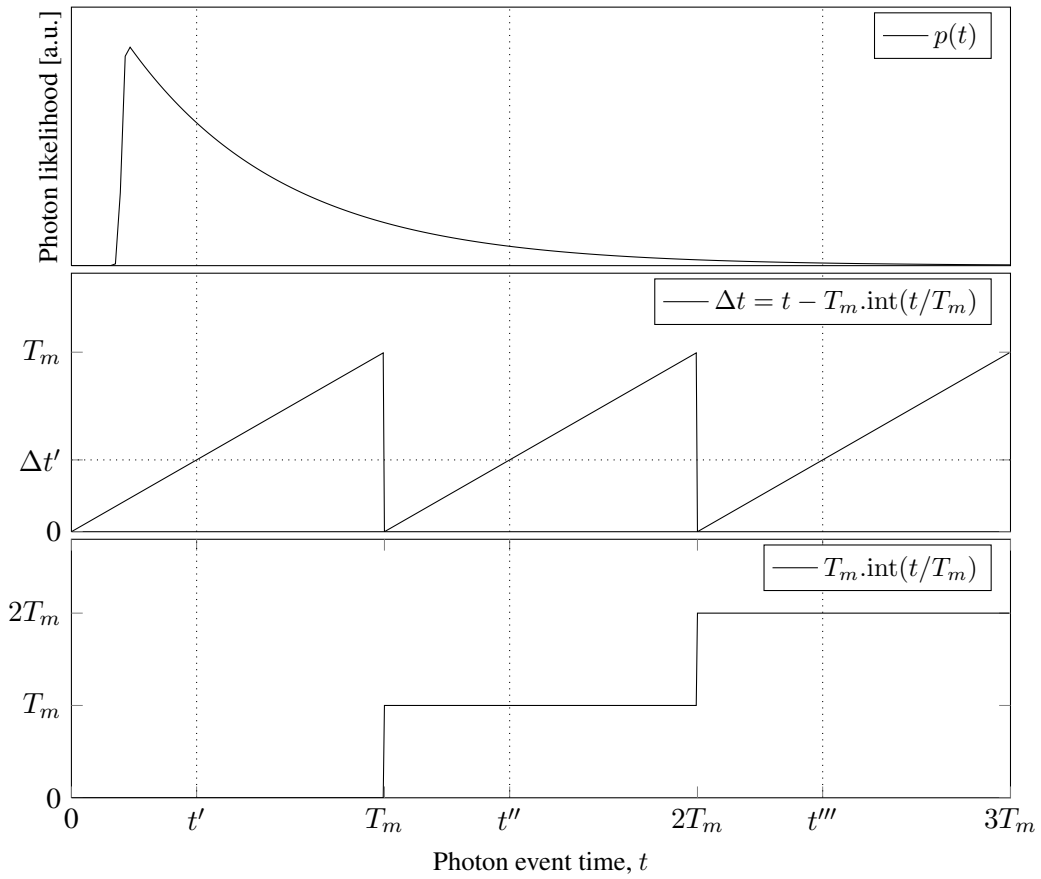


Figure 4.2: *Photon counting on a periodic measurement window:* In the top panel a mono-exponential fluorescence decay of lifetime $T_m/2$ is displayed (having been convolved with a typical instrument response), and in the middle panel the mapping from photon event time t and the recorded event time Δt for photons counted on a periodic measurement window is shown. Notice that although photon counting events occurring at t', $t'' = t' + T_m$, and $t''' = t' + 2T_m$ occur at very different stages of the fluorescence decay, all of them would result in the photon counting event time $\Delta t'$ being recorded at the measurement system due to the periodicity of the measurement window. In the bottom panel, an illustration of the quantity $T_m.\mathrm{int}(t/T_m)$ which restricts the recorded event time Δt to values within the periodic excitation window. Note that in this example an individual fluorescence decay has been isolated in order to consider only the role of the periodic measurement window in the recorded photon counting event time (i.e. the effects of repetitive excitation are neglected here).

z) the actual excitation time being given by $T_m.\mathrm{int}\left((t+u)/T_m\right)$. Observe that, having defined Δt in such a way, there are no assumptions restricting when in the fluorescence

emission process a photon that is subsequently detected had been emitted; a recorded arrival time Δt could be due to a photon emitted a number of repetition periods earlier than the most recent excitation pulse but delayed sufficiently as to be seen in the latest window; or a recorded arrival time may be as a consequence of detected photon that had been emitted due to excitation a period or more beforehand - the likelihood of such situations depending on the fluorescence emission distribution $p(t)$ and the instrument response delay distribution $\Gamma(u)$. Arrival times are recorded only during a measurement interval of duration $T \leq T_m$, and therefore $\Delta t \in [0, T]$. The probability of arrival time Δt for a photon detected in the measurement interval $[0, T]$ on the periodic window, while considering *all* possible emission times t and *all* possible delays u, is given by:

$$p(\Delta t) = \theta(\Delta t)\theta(T - \Delta t) \times \left\{ \frac{w_0}{T} + \frac{(1 - w_0) \int_0^\infty dt du\, p(t)\Gamma(u)\delta\left(\Delta t - t - u + T_m.\mathrm{int}\left(\frac{t+u}{T_m}\right)\right)}{\int_0^T d\Delta t' \int_0^\infty dt du\, p(t)\Gamma(u)\delta\left(\Delta t' - t - u + T_m.\mathrm{int}\left(\frac{t+u}{T_m}\right)\right)} \right\} \tag{4.2}$$

where the emission time t is distributed according to an arbitrary decay signal $p(t)$, the instrument response effects cause a delay u that is distributed according to $\Gamma(u)$, and the quantity $w_0 \in [0, 1]$ represents the contribution of a uniform background to the overall signal. The step function is denoted by $\theta(x)$, with $\theta(x > 0) = 1$ and $\theta(x \leq 0) = 0$, and $\delta(x)$ represents the Delta function which exists only when x is equal to zero, such that $\int dx\ \delta(x)f(x) = f(0)$ for any function $f(x)$. In transforming (as explicitly detailed in Appendix A.1) Eqn. (4.2), the influence that the repetitive nature of the system has on a measured arrival time Δt is instead captured in a summation in the following expression:

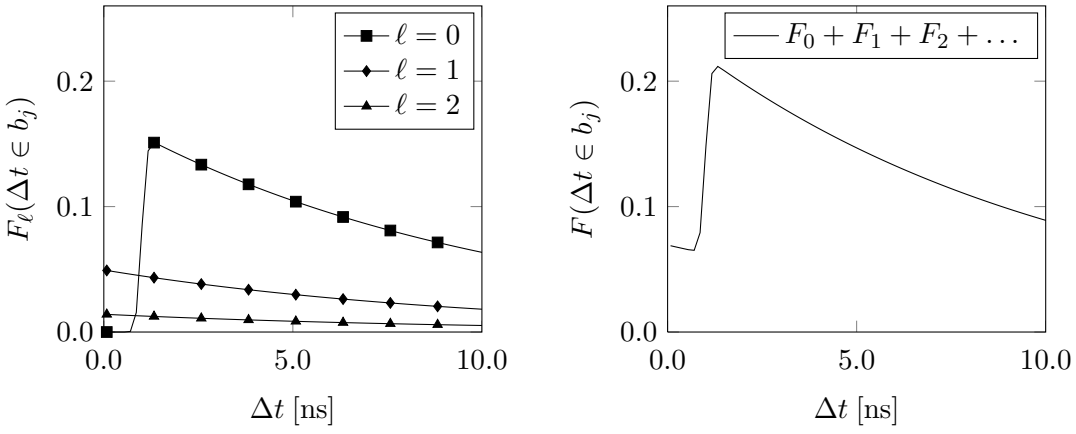
$$p(\Delta t) = \theta(\Delta t)\theta(T - \Delta t) \left\{ \frac{w_0}{T} + \frac{(1 - w_0) \sum_{\ell \geq 0} \int_0^\infty ds\, \Gamma(\ell T_m + \Delta t - s)\, p(s)}{\sum_{\ell \geq 0} \int_0^T d\Delta t' \int_0^\infty ds\, \Gamma(\ell T_m + \Delta t' - s)\, p(s)} \right\} \tag{4.3}$$

The summation has the role of accounting for the history of the signal that may be appar-

ent in the measurement interval, such that when $\ell = 0$ any recorded arrival time is due to a photon emitted and detected in the same measurement interval, when $\ell = 1$ any detected photon had been emitted in the repetition period immediately preceding the measurement interval, when $\ell = 2$ any detected photon emanated from the repetition period before that, and so on (Fig. 4.3). Denoting by $\Lambda(T, T_m)$ the normalisation constant for the term due to fluorescence decay photons, Eqn. (4.3) can be written as,

$$p(\Delta t) = \theta(\Delta t)\theta(T - \Delta t) \left\{ \frac{w_0}{T} + \frac{1 - w_0}{\Lambda(T, T_m)} \int_0^\infty ds\, p(s) \sum_{\ell \geq 0} \Gamma(\ell T_m + \Delta t - s) \right\} \quad (4.4)$$

which remains completely general, incorporating a uniform background proportion w_0, the effects of repetitive excitation (through the summation), an arbitrary instrument response $\Gamma(u)$ and an arbitrary decay signal $p(s)$. The intermediate steps in reaching Eqn. (4.4) are presented in Appendix A.1.1.



(a) Retarded contributions

(b) Bin-likelihood

Figure 4.3: *Repetitive excitation and accumulated fluorescence:* (a) The contribution of the first, second, and third terms in the summation over ℓ of the bin-likelihood $F(\Delta t \in b_j | \tau, b_j^{\mathrm{L}}, b_j^{\mathrm{H}}, \mathcal{I})$, computed for a purely mono-exponential decay of lifetime 10 ns, a system having modulation period 12.5 ns and measurement interval 10.0 ns, and incorporating a single-component instrument response approximation $\Gamma(u)$ having parameter values $\gamma_1 = 1.0$, $u_1 = 1.0$ ns, $\sigma_1 = 0.1$ ns. (b) The bin-likelihood $F(\Delta t \in b_j | \tau, b_j^{\mathrm{L}}, b_j^{\mathrm{H}}, \mathcal{I})$ including the first one hundred terms in the summation over ℓ for the system described in (a).

4.2 Discrete time data

The FLIM data with which this analysis in concerned are actually recorded in discrete time, a photon being determined to have arrived within an interval rather than at a unique time. Accounting for the discrete time nature of such data, the likelihood of photon arrival time Δt being within an interval (i.e. a bin) is required. Defining a bin as being an interval $b = [b^{\mathrm{L}}, b^{\mathrm{H}}] \subseteq [0, T]$ that lies within the measurement window, and adopting the shorthand $p(b) = p(\Delta t \in b)$, the likelihood of a photon arrival time Δt falling in the bin b is given by:

$$p(b) = |b|\frac{w_0}{T} + \frac{1 - w_0}{\Lambda(T, T_m)} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \int_0^{\infty} ds\, p(s) \sum_{\ell \geq 0} \Gamma(\ell T_m + \Delta t - s) \tag{4.5}$$

where $|b| = b^{\mathrm{H}} - b^{\mathrm{L}}$ denotes the width of the interval. Again, notice that Eqn. (4.5) is completely general for a system that generates discretised time-domain data due to the repetitive excitation of a sample (and is subject to the distributions $p(s)$ and $\Gamma(u)$ being normalised).

4.3 Multi-exponential decays

The model developed so far (Eqn. 4.5) describes the time-domain FLIM system, incorporating those effects imposed on the analysis by the design of the experimental system, that is, repetitive excitation and the collection of photon arrival times in discrete time. The introduction of a multi-exponential decay signal $p(s)$ of the form:

$$p(s) = \theta(s)\frac{\sum_{k=1}^{K} \frac{w_k}{\tau_k}\, e^{-s/\tau_k}}{\sum_{k=1}^{K} w_k}, \quad w_k \geq 0,\ \tau_k > 0\ \forall\ k = 1, \ldots, K, \quad \sum_{k=1}^{K} w_k = 1 - w_0 \tag{4.6}$$

into the model (Eqn. 4.5), yields:

$$p(b) = |b|\frac{w_0}{T} + \sum_{k=1}^{K} \frac{w_k}{\tau_k \Lambda(T, T_m)} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \sum_{\ell \geq 0} \int_{0}^{\infty} ds\, \Gamma(\ell T_m + \Delta t - s)\, e^{-s/\tau_k},$$

$$(\boldsymbol{w}, \boldsymbol{\tau}) \in \Omega_K \tag{4.7}$$

where w_k weights the contribution of an exponential decay of lifetime τ_k to the overall signal, the factor $1 - w_0$ has been absorbed due the requirement that $\sum_{k=1}^{K} w_k = 1 - w_0$ and $w_k \in [0, 1]$ for all k. The collection of weight and lifetime parameters are denoted by $\boldsymbol{w} = (w_1, \ldots, w_K)$ and $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_K)$ respectively, and the set Ω_K summarizes all allowable values of the K weights and lifetimes,

$$\Omega_K = \left\{ w_k, \tau_k \middle| w_k \geq 0,\ \tau_k > 0,\ \forall\, k = 1, \ldots, K,\ \sum_{k=1}^{K} w_k \leq 1 \right\}. \tag{4.8}$$

Of course, the accurate analysis of photon arrival time data requires that the effects of the instrument response on the measured photon arrival times be considered, the instrument being assumed to introduce a delay and an uncertainty into the precise time between the physical photon generation and photon detection.

4.4 Analytic instrument response approximation

An approximation comprising a weighted sum of truncated Gaussian distributions is proposed, with the aim that the asymmetry and other artifacts of a real instrument response can be adequately captured:

$$\Gamma(u, \mathcal{I}) = \sum_{i=1}^{I} \gamma_i \frac{e^{-\frac{1}{2}(u-u_i)^2/\sigma_i^2}}{\sigma_i\sqrt{2\pi}} \frac{2\theta[u - \delta_i]}{1 + \mathrm{erf}((u_i - \delta_i)/\sigma_i\sqrt{2})},$$

$$\delta_i, u_i, \sigma_i \geq 0\ \forall i, \quad \gamma_i \in [0, 1]\ \forall i, \quad \sum_{i=1}^{I} \gamma_i = 1 \tag{4.9}$$

where γ_i weights a truncated Gaussian distribution centered about a delay parameter u_i, of width σ_i and having a lower cut-off δ_i, the set $\mathcal{I} = \{\gamma_i, u_i, \sigma_i, \delta_i | i = 1, \ldots, I\}$ summarizes the instrument response parameters, and $\mathrm{erf}(x)$ denotes the error integral $\mathrm{erf}(x) = (2/\sqrt{\pi}) \int_0^x dz\ e^{-z^2}$. In defining this approximation the requirement for a sufficiently good configurable approximation is balanced with the desire that the necessary convolution integrals remain analytic. The inclusion of multiple instrument response components in the approximation provides a means of capturing effects such as signal corruption by the detector, such as detector PMT afterpulsing [103, 104, 60], in the model. The flexibility of the analytic instrument response approximation is demonstrated in Fig. 4.4, as is the influence of an (exaggerated) instrument response function on the bin-likelihood.

The introduction of such an instrument response approximation into the model (Eqn. 4.7) yields (with necessary integration performed analytically as detailed in Appendix A.1.4) the following final expression which describes the likelihood of a photon arrival time Δt being measured in the interval $[b^{\mathrm{L}}, b^{\mathrm{H}}]$ for a system based on repetitive excitation:

$$p(b) = |b| \frac{w_0}{T} + \sum_{k=1}^{K} w_k F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, \mathcal{I}), \qquad (\boldsymbol{w}, \boldsymbol{\tau}) \in \Omega_K, \quad w_0 = 1 - \sum_{k=1}^{K} w_k \tag{4.10}$$

where the fluorescence decay bin-likelihood $F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, \mathcal{I})$ is defined as,

$$F(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) = \sum_i \tilde{F}_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \tag{4.11}$$

and is composed of contributions due to each of the instrument response components,

$$\tilde{F}_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) = \frac{1}{\Lambda(T, T_m)} \tilde{\gamma}_i \sum_{\ell \geq 0} \Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \tag{4.12}$$

The quantity $\Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I})$ is given by the following,

$$\Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) =$$
$$\theta[\ell T_m + b^{\mathrm{L}} - \delta_i] \left\{\chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, b^{\mathrm{L}}, \tau, T_m, \sigma_i, \delta_i, u_i)\right\}$$
$$+ \theta[\ell T_m + b^{\mathrm{H}} - \delta_i] \left\{\chi(\ell, b^{\mathrm{H}}, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i)\right\} \quad (4.13)$$

and,

$$\chi(\ell, t, \tau, \sigma, \delta, u) = \mathrm{erf}\left(\frac{\ell T_m + t - u}{\sigma\sqrt{2}}\right)$$
$$+ e^{-(\ell T_m + t - u)/\tau + \sigma^2/2\tau^2} \left[\mathrm{erf}\left(\frac{(u - \ell T_m - t)\tau + \sigma^2}{\sigma\tau\sqrt{2}}\right) - \mathrm{erf}\left(\frac{(u - \delta)\tau + \sigma^2}{\sigma\tau\sqrt{2}}\right)\right] \quad (4.14)$$

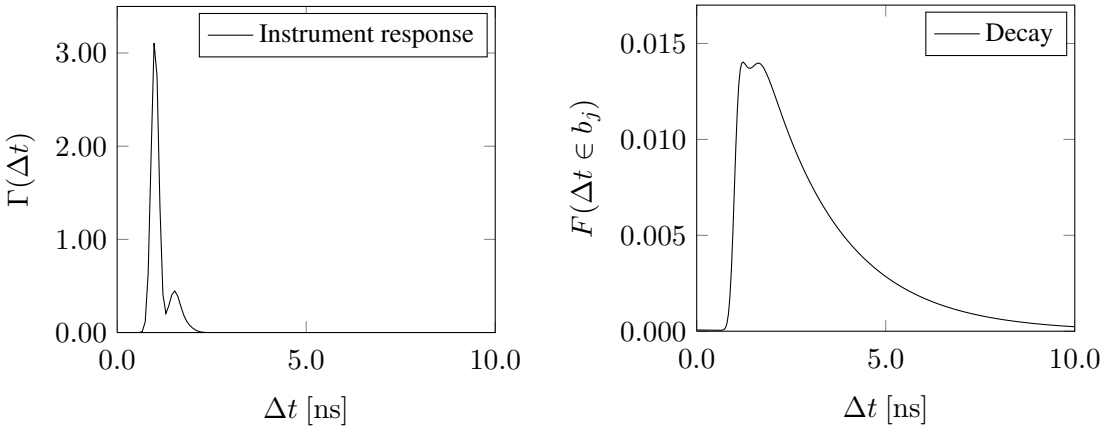
and (for compactness):

$$\tilde{\gamma_i} = \gamma_i \left(1 + \mathrm{erf}((u_i - \delta_i)/\sigma_i\sqrt{2})\right)^{-1} \quad (4.15)$$

It is worthy of note that, although the model (Eqn. (4.10)) incorporates rigorously any history of the signal that may be present in the recorded arrival time data, in practice, for decay lifetimes considerably smaller than the repetition period, the summation over ℓ need only include the first two terms. In the case that a decay lifetime is not significantly smaller than the repetition period, the summation should include more terms as appropriate. The influence of repetitive excitation is illustrated in Fig. 4.3, with the contribution of the first three terms in the summation over ℓ shown in the case that the exponential decay time τ is of the order of the modulation period T_m.

4.5 Why is this model different and what does it offer that others don't?

In this work, considerable effort has been expended to develop a detailed and fully analytic model of a time-domain FLIM system, as described by Eqns. (4.10, 4.11, 4.14, 4.15). The



(a) Multi-component IRF

(b) Decay signal

Figure 4.4: *Instrument response and photon arrival time bin-likelihood:* (a) A three component instrument response approximation $\Gamma(u)$ having parameter values $\gamma_1 = 0.80$, $u_1 = 1.00$ ns, $\sigma_1 = 0.10$ ns, $\gamma_2 = 0.15$, $u_2 = 1.50$ ns, $\sigma_2 = 0.15$ ns, $\gamma_3 = 0.05$, $u_3 = 1.75$ ns, $\sigma_3 = 0.20$ ns. The side lobe is a consequence of the second and third instrument response components, illustrating the flexibility of the analytic approximation employed in this work. (b) The bin-likelihood $F(\Delta t \in b_j|\tau, b_j^{\mathrm{L}}, b_j^{\mathrm{H}}, \mathcal{I})$ computed for a purely mono-exponential decay signal of lifetime 2.0 ns, for a system having modulation period 12.5 ns, a measurement interval of 10.0 ns partitioned into 256 bins of equal width, perturbed by the instrument response $\Gamma(u)$ as detailed above. Notice, particularly around the peak of the bin-likelihood, that the influence of the side lobe of the instrument response approximation is clearly apparent.

developed model incorporates, of course, a multi-exponential fluorescence decay process and also accounts formally for at least those features that may have a considerable bearing on parameter estimates when the model is used for the quantification of a fluorescence decay process. On first inspection of the analytic FLIM system model it is often remarked that the most striking difference is that that the equations which define the FLIM system model appear to be much more complicated than those familiar from the traditional fitting techniques often used for fluorescence decay analysis. The main reasons for there being many more variables and an apparent complexity in Eqns. (4.10, 4.11, 4.14, 4.15) that do not exist in the familiar approaches to fluorescence decay analysis are:

- *A more complete approach:*
  The FLIM system model has been developed to more faithfully represent the time-resolved FLIM system by formally incorporating the effects of repetitive excitation;

this requires a variable to describe the repetition period (T_m) and a more sophisticated definition of the photon arrival time $\Delta t = t + u - T_m.\mathrm{int}\left((t+u)/T_m\right)$ that avoids the assumption that any detected fluorescence decay photon had been a result of the excitation at the beginning of the measurement window.

- *A fully analytic approach:*
  The FLIM system model has been developed to be fully analytic thereby requiring additional variables to realise an analytic approximation to the instrument response; for the chosen approximation in this work the variables $\mathcal{I} = \{\gamma_i, u_i, \sigma_i, \delta_i | i = 1, \ldots, I\}$ describe the weighted sum of set of truncated Gaussian distributions. Much of the apparent complexity of the FLIM system model is a consequence of it being fully analytic and as such the convolution integrals in Eqn. (4.7) have been performed analytically; for the chosen instrument response approximation and a multi-exponential decay signal this results in the error function integrals that appear in Eqn. (4.14). Typically, in implementing the traditional fitting methods convolution of a measured instrument response with a proposed decay signal is performed numerically and therefore there is no requirement for the computation of convolution integrals akin to those in Eqn. (4.7).

- *A probabilistic approach:*
  As the intended application of the FLIM system model is within the Bayesian framework, careful attention has been paid throughout its development to ensure that all quantities that comprise the FLIM system model are appropriately normalised. Particularly, as only those photon arrival times that can be measured are considered in the analysis, the inclusion of a variable to describe the measurement window (T) is required.

In the remainder of this section a critical comparison of the FLIM system model with the models used in other approaches to fluorescence decay analysis is undertaken, particularly the additional variables and the apparent complexity of the FLIM system model are justified by demonstrating theoretically what the FLIM system model offers over the other approaches to FLIM analysis that were discussed in Chapter 2.

In the direct fitting approach to time-domain FLIM data a fluorescence decay model is fitted directly to the photon counting histogram, the optimal fit being determined according to a goodness-of-fit parameter which somehow quantifies the distance between the fit

and the measured photon counting histogram. The key ingredients of the direct fitting approach are usually a measured instrument response function and a fluorescence decay model which is typically of the form,

$$I(t) = Z + \sum_{\ell=1}^{L} A_\ell e^{-t/\tau_\ell} \tag{4.16}$$

where $I(t)$ represents the fluorescence intensity (the photon count) at time t, Z represents the constant background level, and each decay component is described by an initial intensity A_ℓ and a decay lifetime τ_ℓ. In the direct fitting approach, the optimal fit is determined by finding those parameter values that minimise the goodness-of-fit parameter; in performing such an analysis, the necessary deconvolution with a (usually) measured instrument response is typically performed numerically.

Clearly, as both the direct fitting approach and the FLIM system model are designed to be used for the quantification of a fluorescence decay process it is no surprise that they do have similarities; although the conventional fluorescence decay model of Eqn. (4.16) operates directly in units of photon counts and is applied directly to the photon count histogram whereas the FLIM system model is probabilistic and applied for each counted photon, the analogues between the fitting decay model (Eqn. (4.16)) and its Bayesian counterpart (Eqn. (4.6)) are easily identifiable. However, the differences between the two approaches are more numerous, as explored below.

As a measured instrument response is used in the direct fitting approach and the required convolution operations are performed numerically there are no counterpart variables in the direct fitting approach to those in Eqns. (4.10, 4.11, 4.14, 4.15) that describe the analytic instrument response approximation of the FLIM system model. While it could be argued that the integrity of the effects of the FLIM system in the data might be better preserved by using the measured instrument response and performing numerical integration for computation of the integrals in Eqn. (4.14), it should also be noted that the measurement of an instrument response may not be straightforward (nor even possible in some circumstances), and even then that the measurement itself may not be perfect. Indeed, it is not always an accurate assumption that a reliable measured instrument response is available for use in decay data analysis. It is also worthwhile remarking that, although it is usually the quantification of a fluorescence decay process that is of interest, the fluorescence decay variables hold no greater status than any of the other variables in

the FLIM system model. The presence of those variables in Eqns. (4.10, 4.11, 4.14, 4.15) that describe the instrument response approximation, and those defining the measurement interval and repetition period, should certainly not be viewed as merely an additional complication in the theory; the very presence of these variables enables properties of the FLIM system to be approximated or potentially even optimised from the decay data. In this work those variables that describe analytically an instrument response approximation have been exploited in order that the optimal instrument response approximation can be estimated from time resolved data alone (without any recall to the measured instrument response even if such is available).

In contrast to the developed FLIM system model, where the multi-exponential decay model and the analytic instrument response are combined in a framework (Eqn. (4.4)) that rigourously respects repetitive excitation in time-domain FLIM to yield the likelihood of a photon being counted at *any* time within the measurement window, in the direct fitting approach the decay model and instrument response are combined without repetitive excitation being accounted for. In neglecting the effects of repetitive excitation in the analysis, the direct fitting approach that combines a fluorescence decay model of the form of Eqn. (4.16) and a measured instrument response cannot account for any fluorescence decay photons that may have been counted early in the measurement window (i.e. fluorescence be non-negligible towards the end of an excitation period). In effect, the fact that the counted photons actually emanate from a long train of fluorescence decays extending back in time over many excitations is lost and (although it will always be formally incorrect) the analysis will only yield reliable decay parameter estimates when the fluorescence is certain to be negligible towards the end of an excitation period (i.e. the lifetime of the slowest fluorescence decay component is significantly less than the excitation period). The variables T and T_m that do exist in the FLIM system model relations (Eqns. (4.10, 4.11, 4.14, 4.15)) but not in the direct fitting approach could be viewed as a replacement for the need of ad hoc rules that must be obeyed for an analysis to be valid. The addition of these variables is critical in this work and important in providing a more complete picture of the FLIM system than that of the assumptions on which conventional analyses are based. Certainly, should the technology be available and there be a desire to perform an experiment that does not meet the usually accepted restrictions (i.e. decay lifetime and excitation period), then it would be a shame if that experiment not be progressed due to an incomplete approach to data analysis.

It is easy to see that it would not be possible for the direct fitting approach, which does not account for repetitive excitation, to describe decay data arising from a fluorescence decay such as that shown in Fig. 4.5, in which a baseline fluorescence accumulates over time. One could decide to apply the direct fitting approach incorporating a decay model of the form of Eqn. (4.16) only to data following the rise of the transient (for example, in Fig. 4.3 for photons that would be counted after about 2.0 ns in the measurement window); this is however an unsatisfactory workaround that would almost certainly yield incorrect background and initial intensity estimates. Although this may not be of significant importance for a mono-exponential analysis as these parameters are not usually of much interest, should an estimate of the interacting fraction (i.e. A_2/A_1) be required for a FRET study, then a bi-exponential analysis of data incorporating a fluorescence component similar to that of Fig. 4.5 using the direct fitting approach would almost certainly yield incorrect initial amplitudes and an incorrect FRET interacting fraction. In [67] the influence of repetitive excitation was accounted for theoretically for use in the direct fitting approach to yield a modified fluorescence decay model of the form (variable names and index labels have been changed),

$$I(t) = Z + I_0 \sum_{\ell=1}^{L} R_\ell \alpha_\ell e^{-t/\tau_\ell}, \qquad R_\ell = 1 + \frac{1}{e^{T_m/\tau_\ell} - 1}, \qquad \sum_\ell \alpha = 1 \tag{4.17}$$

where α_ℓ weights the contribution of the $\ell\mathrm{th}$ decay component to the overall fluorescence intensity and the factor R_ℓ accounts for the apparent increased initial amplitude as a consequence of repetitive excitation. Although such a modified decay model does incorporate the influence of repetitive excitation and should yield correct initial amplitudes when used as the basis of the direct fitting approach if applied only to the *right* data in the measurement window, unless it were applied in a framework that somehow captures that the instrument does not respond instantly and that fluorescence photons from a previous repetition period will be counted in the measurement window, it cannot be successfully applied to the entire measurement window (i.e. all of the recorded data).

In the phasor approach to time-resolved FLIM data, as presented in [70], a fluorescence intensity model $I(t)$ is incorporated in defining the phasor coordinates,

$$g(\omega) = \frac{\int_0^\infty dt\ I(t)\cos(\omega t)}{\int_0^\infty dt\ I(t)}, \qquad s(\omega) = \frac{\int_0^\infty dt\ I(t)\sin(\omega t)}{\int_0^\infty dt\ I(t)} \tag{4.18}$$

where ω is the laser repetition angular frequency (i.e. $\omega = 2\pi/T_m$). Obviously, should the fluorescence intensity model be chosen to be of the form of Eqn. (4.16) then the arguments regarding the limitations of the approach also apply. The mono-exponential phasor analysis presented in [70] incorporates a model of the form $I(t) = Ae^{-t/\tau}$ (notice that a constant background component is not allowed for). It should also be observed that (unless somewhere hidden in the fluorescence decay model $I(t)$) the effects of the instrument response of the time resolved FLIM data is not considered in the phasor approach as stated. As will be shown in Chapter 8 neglecting the influence of the instrument response in data analysis has considerable risks.

In summary, although in *some* circumstances one can obtain good fluorescence decay parameter estimates using, say, the direct fitting approach based on a model of the form of Eqn. (4.16) this is certainly not true in *all* circumstances. It is not surprising that, no matter how good a particular analysis technique may be (e.g. LS, ML, Bayesian analysis), or even were a *perfect* analysis method to exist, without a sufficiently detailed model of the time domain FLIM system then accurate fluorescence decay parameter estimates should not be taken for granted.

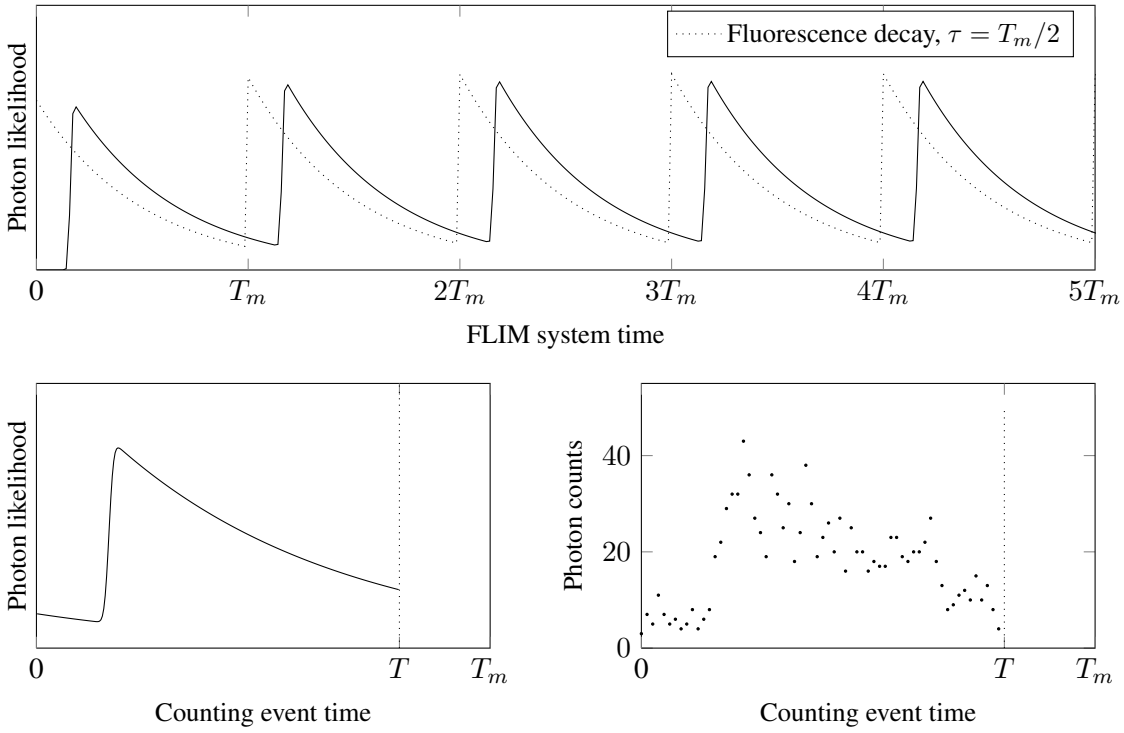


Figure 4.5: *Accumulating fluorescence; repetitive excitation, periodic fluorescence decay and photon counting with a slow decay:* For a simulated mono-exponential decay of lifetime $T_m/2$, a typical instrument response, and a FLIM system repetition period of T_m, in the uppermost panel the initial sequence of periodic fluorescence decays (in FLIM system time) as a consequence of repetitive excitation; notice particularly the increase in photon likelihood at subsequent decay peaks as fluorescence accumulates with passing FLIM system time. In the bottom left panel, the fluorescence decay photon likelihood accounting for repetitive excitation of the FLIM system within the measurement interval $[0, T]$; observe again the accumulated fluorescence elevating the photon likelihood over the entire measurement window and also the history of decays due to previous excitations at the beginning of the measurement window. Typical photon count data having about 1000 total photon counts is shown in the bottom right panel.

Chapter 5

Bayesian analysis applied to time resolved FLIM

In this chapter, the Bayesian framework is applied to the FLIM forward model developed above (Eqn. (4.10)) with the purpose of quantifying probabilistically the source of some fluorescence decay data and also to provide answers to (the usually unasked) questions regarding the nature of both the fluorescence decay and of the FLIM instrument. In particular, the following are explored:

- ***Parameter estimation***
  What is the probability of a set of fluorescence decay model parameter values given the data, the decay model, and the parameterized instrument response approximation?
- ***Model selection***
  How many exponential decay components are likely to have yielded the decay data?
- ***Instrument response determination***
  What is the most likely form of the instrument response function given the *fluorescence decay* data?

The following notation is introduced and shall be used throughout the remainder of this document:

- **The data**
  The data $D = \{(b_j, c_j)|j = 1, \ldots, M\}$ comprise a set of M bin-count pairs (b_j, c_j) where b_j represents the jth bin (i.e. interval $b_j = [b_j^{\mathrm{L}}, b_j^{\mathrm{H}}] \in [0, T]$) and c_j is the number of photons recorded as having been counted into that bin. It is required that none of the bins overlap (i.e. $b_j \cap b_k = \emptyset,\ \forall j \neq k$) and that their union forms the measurement interval (i.e. $\cup_{j=1}^{M} b_j = [0, T]$).

- **The decay model**
  The fluorescence decay model is denoted by $\mathcal{H}_K$, where K exponential decays of lifetime τ_k contribute to the overall decay according to weight w_k as defined by Eqn. (4.6). The notation $\boldsymbol{w}_K = (w_1, \ldots, w_K)$ and $\boldsymbol{\tau}_K = (\tau_1, \ldots, \tau_K)$ shall also be used.

- **The parameterized instrument response**
  The characteristics of the FLIM equipment are denoted by $\mathcal{I} = \{\gamma_i, u_i, \sigma_i, \delta_i|i = 1, \ldots, I\}$, which defines an instrument response approximation composed of I weighted, truncated Gaussian distributions, as defined by Eqn. (4.9).

The Bayesian analysis developed in this section intentionally does not explicitly incorporate any particular form for the required prior distributions, in order that the expressions developed remain general and can be used as a starting point and be suitably updated on making specific the choice of prior distribution.

5.1 Decay parameter estimation

Most commonly, it is the fluorescence decay parameter value estimates that are of greatest interest to the experimentalist, with both the decay model $\mathcal{H}_K$ and the instrument parameters $\mathcal{I}$ being fixed having been either determined or assumed. The posterior distribution gives the likelihood of fluorescence decay parameter values $(\boldsymbol{w}_K, \boldsymbol{\tau}_K)$ given the data D, the instrument response characteristics $\mathcal{I}$, and the decay model $\mathcal{H}_K$:

$$
\begin{aligned}
&p(\boldsymbol{w}_K, \boldsymbol{\tau}_K | D, \mathcal{H}_K, \boldsymbol{\alpha}_K, \mathcal{I}) \\
&\quad = \frac{p(\boldsymbol{w}_K, \boldsymbol{\tau}_K | \mathcal{H}_K, \boldsymbol{\alpha}_K) p(D | \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{H}_K, \mathcal{I})}{\int d\boldsymbol{w}'_K \int d\boldsymbol{\tau}'_K \; p(\boldsymbol{w}'_K, \boldsymbol{\tau}'_K | \mathcal{H}_K, \boldsymbol{\alpha}_K) p(D | \boldsymbol{w}'_K, \boldsymbol{\tau}'_K, \mathcal{H}_K, \mathcal{I})} \\
&\quad = \frac{p(\boldsymbol{w}_K, \boldsymbol{\tau}_K | \mathcal{H}_K, \boldsymbol{\alpha}_K) \prod_{j=1}^{M} p(b_j | \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{H}_K, \mathcal{I})^{c_j}}{\int d\boldsymbol{w}'_K \int d\boldsymbol{\tau}'_K \; p(\boldsymbol{w}'_K, \boldsymbol{\tau}'_K | \mathcal{H}_K, \boldsymbol{\alpha}_K) \prod_{j=1}^{M} p(b_j | \boldsymbol{w}'_K, \boldsymbol{\tau}'_K, \mathcal{H}_K, \mathcal{I})^{c_j}}
\end{aligned}
\tag{5.1}
$$

Of course, as any parameter estimates $(\boldsymbol{w}_K, \boldsymbol{\tau}_K)$ derived from Eqn. (5.1) are conditioned not only on the data D but also on the decay model $\mathcal{H}_K$ (and its hyperparameters $\boldsymbol{\alpha}_K$) and instrument response characterization $\mathcal{I}$, it should be expected that they be most reasonable when the appropriate $\mathcal{H}_K$ and $\mathcal{I}$ have been employed throughout the analysis. The Bayesian approach to determination of the decay model and instrument response characteristics is developed in the following sections.

5.2 Decay model selection

Here, in order to address such questions as "*Should I be using mono-exponential or bi-exponential analysis?*", the Bayesian framework is employed to determine the most applicable *fluorescence decay* model $\mathcal{H}_K$ given the data, where it is assumed in this analysis that the instrument response parametrization $\mathcal{I}$ has been appropriately determined. In this work, the most probable decay model and the most probable hyperparameter(s) are determined together, their posterior distribution offering a quantitative measure of the likelihood of the recorded data being due to the model $\mathcal{H}_K$ its hyperparameter(s) $\boldsymbol{\alpha}_K$, as given by,

$$
\begin{aligned}
&p(\mathcal{H}_K, \boldsymbol{\alpha}_K | D, \mathcal{I}) \\
&\quad = \frac{p(\mathcal{H}_K, \boldsymbol{\alpha}_K) \int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K | \mathcal{H}_K, \boldsymbol{\alpha}_K) p(D | \mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I})}{\sum_{K'} \int d\boldsymbol{\alpha}_{K'} \; p(\mathcal{H}_{K'}, \boldsymbol{\alpha}_{K'}) \int d\boldsymbol{w}_{K'} d\boldsymbol{\tau}_{K'} \; p(\boldsymbol{w}_{K'}, \boldsymbol{\tau}_{K'} | \mathcal{H}_{K'}, \boldsymbol{\alpha}_{K'}) p(D | \mathcal{H}_{K'}, \boldsymbol{w}_{K'}, \boldsymbol{\tau}_{K'}, \mathcal{I})}
\end{aligned}
\tag{5.2}
$$

where $p(\mathcal{H}_K, \boldsymbol{\alpha}_K)$ is the prior probability of decay model $\mathcal{H}_K$ and its hyperparameter(s)

$\boldsymbol{\alpha}_K$, and in the denominator the summation is over all candidate decay models. Using Eqn. (5.2), the relative likelihood of candidate decay models can be obtained by the pairwise comparison of the evidence for those different models,

$$\begin{aligned}
&\frac{p(\mathcal{H}_K, \boldsymbol{\alpha}_K|D, \mathcal{I})}{p(\mathcal{H}_{K'}, \boldsymbol{\alpha}_{K'}|D, \mathcal{I})} \\
&\quad = \frac{p(\mathcal{H}_K, \boldsymbol{\alpha}_K)}{p(\mathcal{H}_{K'}, \boldsymbol{\alpha}_{K'})} \frac{\int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K|\mathcal{H}_K, \boldsymbol{\alpha}_K) p(D|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I})}{\int d\boldsymbol{w}_{K'} d\boldsymbol{\tau}_{K'} \; p(\boldsymbol{w}_{K'}, \boldsymbol{\tau}_{K'}|\mathcal{H}_{K'}, \boldsymbol{\alpha}_{K'}) p(D|\mathcal{H}_{K'}, \boldsymbol{w}_{K'}, \boldsymbol{\tau}_{K'}, \mathcal{I})}
\end{aligned} \tag{5.3}$$

Alternatively, the most probable decay model and its hyperparameters conditioned on the data can be sought by finding the maximum of Eqn. (5.2). Since $p(\mathcal{H}_K, \boldsymbol{\alpha}_K) = p(\mathcal{H}_K) p(\boldsymbol{\alpha}_K|\mathcal{H}_K)$, the most probable decay model $\mathcal{H}_K^\star$ and its optimal hyperparameters $\boldsymbol{\alpha}_K^\star$ are given by,

$$\begin{aligned}
(\mathcal{H}_K^\star, \boldsymbol{\alpha}_K^\star) &= \underset{\mathcal{H}_K, \boldsymbol{\alpha}_K}{\operatorname{argmax}} \left[p(\mathcal{H}_K) p(\boldsymbol{\alpha}_K|\mathcal{H}_K) Z(\mathcal{H}_K, \boldsymbol{\alpha}_K, \mathcal{I}) \right] \\
&= \underset{\mathcal{H}_K}{\operatorname{argmax}} \left[p(\mathcal{H}_K) \max_{\boldsymbol{\alpha}_K} \left[p(\boldsymbol{\alpha}_K|\mathcal{H}_K) Z(\mathcal{H}_K, \boldsymbol{\alpha}_K, \mathcal{I}) \right] \right] \\
&= \underset{\mathcal{H}_K}{\operatorname{argmax}} \left[p(\mathcal{H}_K) p(\boldsymbol{\alpha}_K^\star|\mathcal{H}_K) Z(\mathcal{H}_K, \boldsymbol{\alpha}_K^\star, \mathcal{I}) \right],
\end{aligned} \tag{5.4}$$

where the integral $Z(\mathcal{H}_K, \boldsymbol{\alpha}_K, \mathcal{I})$ is given by,

$$Z(\mathcal{H}_K, \boldsymbol{\alpha}_K, \mathcal{I}) = \int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K|\mathcal{H}_K, \boldsymbol{\alpha}_K) p(D|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}), \tag{5.5}$$

the optimal hyperparameter(s) are given by,

$$\boldsymbol{\alpha}_K^\star = \underset{\boldsymbol{\alpha}_K}{\operatorname{argmax}} \left[p(\boldsymbol{\alpha}_K|\mathcal{H}_K) Z(\mathcal{H}_K, \boldsymbol{\alpha}_K, \mathcal{I}) \right], \tag{5.6}$$

and $p(\boldsymbol{\alpha}_K|\mathcal{H}_K)$ (the so-called hyperprior) is the prior distribution of the hyperparameter(s) $\boldsymbol{\alpha}_K$ for the decay model $\mathcal{H}_K$. The expressions for fluorescence decay model selection are developed further in Appendix A.2 for the decay model and prior specific

to this work. In this work, the integrals that must be determined for model selection (Eqns. 5.2, 5.3, 5.4, 5.6) are computed using a Gaussian approximation, as detailed in Appendix A.2.1.

5.3 Instrument response determination

The measurement of an instrument response can sometimes be a difficult practical problem. This section develops the Bayesian determination of the instrument response approximation parameters from the fluorescence decay data itself, obviating the need for any separate experiment with the sole purpose of measuring the instrument response. Denoting by $\mathcal{I} = \{u_i, \sigma_i, \delta_i | i = 1, \ldots, I\}$ the parameter values of the instrument response approximation $\Gamma(u)$ (Eqn. (4.9)), the posterior distribution of the instrument response parameter values is given by:

$$
\begin{aligned}
& p(\mathcal{I}|D) \\
& \quad = \frac{p(\mathcal{I}) \sum_{\mathcal{H}_K} p(\mathcal{H}_K) \int d\boldsymbol{w}_K \int d\boldsymbol{\tau}_K \, p(\boldsymbol{w}_K, \boldsymbol{\tau}_K) p(D|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I})}{\sum_{\mathcal{H}_K} p(\mathcal{H}_K) \int d\boldsymbol{w}_K \int d\boldsymbol{\tau}_K \, p(\boldsymbol{w}_K, \boldsymbol{\tau}_K) p(D|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I})} \\
& \quad = \frac{p(\mathcal{I}) \sum_{\mathcal{H}_K} p(\mathcal{H}_K) \int d\boldsymbol{w}_K \int d\boldsymbol{\tau}_K \, p(\boldsymbol{w}_K, \boldsymbol{\tau}_K) \prod_{j=1}^{M} p(b_j|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I})^{c_j}}{\int d\mathcal{I}' \, p(\mathcal{I}') \sum_{\mathcal{H}_K} p(\mathcal{H}_K) \int d\boldsymbol{w}_K \int d\boldsymbol{\tau}_K \, p(\boldsymbol{w}_K, \boldsymbol{\tau}_K) \prod_{j=1}^{M} p(b_j|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}')^{c_j}}
\end{aligned}
\tag{5.7}
$$

Of course, should the data be acquired using a fluorophore known to exhibit a purely mono-exponential decay, the above simplifies to yield:

$$
p(\mathcal{I}|D, \mathcal{H}_1) = \frac{p(\mathcal{I}) \int d\boldsymbol{w}_1 \int d\boldsymbol{\tau}_1 \, p(\boldsymbol{w}_1, \boldsymbol{\tau}_1) \prod_{j=1}^{M} p(b_j|\boldsymbol{w}_1, \boldsymbol{\tau}_1, \mathcal{I})^{c_j}}{\int d\mathcal{I}' \, p(\mathcal{I}') \int d\boldsymbol{w}_1 \int d\boldsymbol{\tau}_1 \, p(\boldsymbol{w}_1, \boldsymbol{\tau}_1) \prod_{j=1}^{M} p(b_j|\boldsymbol{w}_1, \boldsymbol{\tau}_1, \mathcal{I}')^{c_j}} \tag{5.8}
$$

Additionally, it is also possible to estimate the instrument response parameters at the same time as estimating the fluorescence decay parameters (as has been previously implemented and used to date). Again, under the assumption of a purely mono-exponential decay, the posterior to be calculated simplifies yet further:

$$p(\boldsymbol{w}_1, \boldsymbol{\tau}_1, \mathcal{I}|D, \mathcal{H}_1) = \frac{p(\mathcal{I})p(\boldsymbol{w}_1, \boldsymbol{\tau}_1)\prod_{j=1}^{M} p(b_j|\boldsymbol{w}_1, \boldsymbol{\tau}_1, \mathcal{I})^{c_j}}{\int d\mathcal{I}'\; p(\mathcal{I}') \int d\boldsymbol{w}_1' \int d\boldsymbol{\tau}_1'\; p(\boldsymbol{w}_1', \boldsymbol{\tau}_1') \prod_{j=1}^{M} p(b_j|\boldsymbol{w}_1', \boldsymbol{\tau}_1', \mathcal{I}')^{c_j}} \tag{5.9}$$

Part III

Results

The desire for more accurate mono-exponential and bi-exponential fluorescence decay parameter estimates with fewer photon counts provoked much of the work that has lead to the fully analytic FLIM system model and the Bayesian decay analysis algorithms as presented in Part II, and to the results presented here for mono-exponential and bi-exponential fluorescence decay analysis (Chapters 6 & 7 respectively). However, Bayesian analysis using the fully analytic FLIM system model is able to offer considerably more than merely being an additional (and, as the reader may be persuaded in the coming chapters, enhanced) decay analysis technique. The incorporation of parameters to describe an approximation to the system IRF into the fully analytic FLIM system model makes it possible to estimate the system IRF given the decay data alone, as is realised by the Bayesian SID (simultaneous instrument and decay) algorithm. Additionally, the hierarchical nature of Bayesian analysis readily permits the development of model selection analysis; in this work a Bayesian decay model selection algorithm that quantifies the relative likelihoods of different decay orders (e.g. background only, mono-exponential and bi-exponential) from the decay data is presented. Before discussing the performance of the different Bayesian algorithms it is useful to first clarify the exact purpose of each algorithm, which of the FLIM system model parameters the algorithm is used to estimate, and which parameter values must be provided for a successful analysis. The fully analytic FLIM system model developed in Part II contains variables that describe a multi-exponential decay and variables that describe an analytic approximation to the instrument response. As summarised below, each of the Bayesian analysis algorithms has been developed for the purpose of estimation of one or more of these groups of variables, and requires that the values of all of the other variables be provided, as summarised below:

- *Bayesian decay analysis for fluorescence decay parameter estimation:*
  Estimation of the optimal parameters that describe a fluorescence decay (Eqn. 4.6), i.e. $\{w_k, \tau_k | k = 1, \ldots, K\}$, where w_k weights the contribution of an exponential decay of lifetime τ_k to the overall signal having K components. For a mono-exponential decay the Bayesian decay parameters are $\{w_1, \tau_1\}$ and for a bi-exponential decay $\{w_1, \tau_1, w_2, \tau_2\}$ and their conventional direct fitting counterparts are $\{A_1, \tau_1\}$ and $\{A_1, \tau_1, A_2, \tau_2\}$ respectively. The parameters which describe the IRF approximation are required and should be optimal for effective Bayesian decay analysis and the decay order K is set by the user.

- *Bayesian SID analysis for fluorescence decay and IRF approximation parameter estimation:*
  Simultaneous estimation of the optimal fluorescence decay parameter values and those parameters which describe the analytic IRF approximation (Eqn. 4.9), i.e. $\mathcal{I} = \{\gamma_i, u_i, \sigma_i, \delta_i | i = 1, \ldots, I\}$, where γ_i weights the contribution of a truncated Gaussian distribution centered about a delay parameter u_i, of width σ_i and having a lower cut-off δ_i, to an instrument response approximation comprised of I such distributions. Bayesian SID analysis can be applied for either a mono-exponential or a bi-exponential decay model.

- *Bayesian decay model selection for fluorescence decay model order estimation:*
  Determination of the optimal decay model $\mathcal{H}_K$ by estimation of the relative likelihood of the decay data being due to a mono-exponential ($\mathcal{H}_1$) or bi-exponential ($\mathcal{H}_2$) decay process. In applying the Bayesian decay model selection algorithm the optimal mono-exponential and bi-exponential fluorescence decay parameters are also estimated; the algorithm requires that the optimal IRF approximation is provided.

The developed Bayesian mono-exponential and bi-exponential decay analysis algorithms were tested against ML and LS, for mono-exponential data against phasor analysis, and for bi-exponential data against global analysis, all approaches operating directly on the accumulated histogram. The ML estimation routines were implemented as described in [82], and are based on the modified Levenberg-Marquardt (MLM) algorithm. The LS implementation is also based upon the MLM algorithm and is described in [67, 87]. Phasor analysis (e.g. [90, 105]) was implemented incorporating automated background level estimation at each pixel (using data before the rise of the transient), with Fourier components computed using modulation frequency $\omega = 2\pi/m$ set according to the number of time bins m that form the window of data to be analysed, and with the estimated background level being subtracted from the photon count at each bin of the data window. The average of the phase and modulation lifetime values is reported. The bi-exponential global analysis algorithm was implemented to operate under assumption that the decay lifetimes are invariant over each analysed image. As the repetitive nature of TCSPC excitation is not accounted for in any of the established analysis techniques used for comparison, in order to avoid their potential effects the ML, LS, and phasor analysis routines all consider only a window of the collected data that excludes time points before the rise of the tran-

sient has occurred. Automated selection of the valid data window is discussed in [67]. The Bayesian analysis routines operate on all of the available data, except for small portions at the start and end of the transient, potentially corrupted by consequences of dithering associated with the time-amplitude converter in the TCSPC electronics [60].

The Bayesian analysis algorithms were implemented in the C programming language and incorporated into the TRI2 (Time Resolved Imaging 2) software package [67, 87], as illustrated in Fig. 5.1. A significant amount of effort has been expended in developing a "user-friendly" Bayesian implementation that is accessible without having knowledge of the FLIM system model or the Bayesian analysis of Chapters 4 and 5 respectively, or even of Bayesian analysis in general (Chapter 3). At the time of the initial integration of the Bayesian algorithms into TRI2, efforts were made such that their operation should be as similar as possible, from the point of view of user-interaction, to that of the existing LS analysis algorithms, such that a user familiar with the LS routines should be able to use the Bayesian algorithms without facing a steep learning curve. Indeed, one of the key *results* of the work that has culminated in this thesis is the realisation of an accessible implementation, as discussed in Appendix C. TRI2 also provides ML, LS, and phasor analysis fitting routines for the analysis of time-resolved data and is widely used in the laboratory in the Randall Division of Cell & Molecular Biophysics at King's College London and at the Gray Institute for Radiation Oncology and Biology at the University of Oxford (where most of the work that has culminated in this thesis has been conducted).

The analysis algorithms were tested with simulated data of the expected signal from a model of the TCSPC system which included a fixed background and the effect of repetitive excitation [67] at 40 MHz; simulated Poisson noise was added to the convolution of a raw fluorescence signal and a Gaussian excitation pulse to simulate realistic photon count transients. Additionally, the algorithms were also compared with experimental biological data, as detailed in Appendix B.2.

The performance of the developed Bayesian mono-exponential and bi-exponential decay analysis algorithms is compared with established analysis techniques in Chapter 6 & 7 respectively. In comparing the parameter estimates of the different techniques both their accuracy and precision are presented; the closeness to the true parameter value of the average of a distribution of estimates providing a measure of accuracy of the parameter estimates with the uncertainty being quantified by the standard deviation of the distribution of the estimates. All of the Bayesian fluorescence decay analysis results presented in

Chapter 6 & 7 were obtained using the optimal IRF approximation as determined by application of the Bayesian SID algorithm to representative very high total photon count data. The results of the application of the Bayesian SID algorithm are presented in Chapter 8. As Bayesian SID analysis is used to estimate the decay parameters *and* an approximation to the IRF *given* the decay data, in a typical analysis its application would usually precede a Bayesian decay analysis in order to provide an optimal approximation to the IRF for use by the Bayesian decay analysis algorithms. Bayesian decay model selection could reasonably be used prior to or following a decay analysis, though as it builds upon Bayesian decay analysis it is presented in Chapter 9. In order to present a digestible account of the performance of the Bayesian analysis algorithms the results presented and the examples explored in the following chapters are focused both towards areas where the Bayesian algorithms are likely to find useful purpose and offer an advantage over other analysis techniques and to those where existing techniques either match or better the overall performance of the Bayesian algorithms.

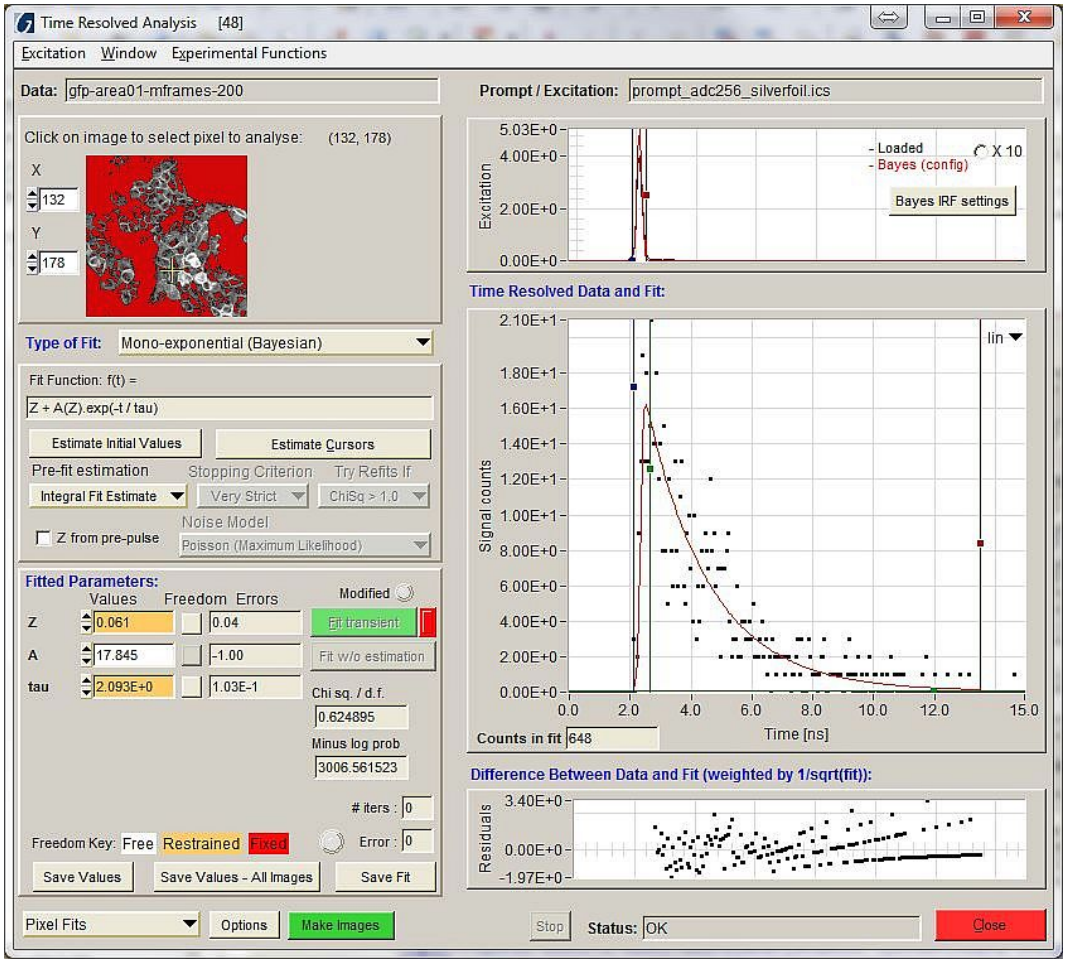


Figure 5.1: *Bayesian analysis of time-resolved data in TRI2:* The results of mono-exponential Bayesian analysis of an image pixel containing 648 photon counts; notice that the mono-exponential Bayesian parameter estimates and their uncertainty, the fitted fluorescence decay and its residuals, and also the Bayesian approximation to the instrument response and the loaded measured instrument response, are all displayed to the user.

Chapter 6

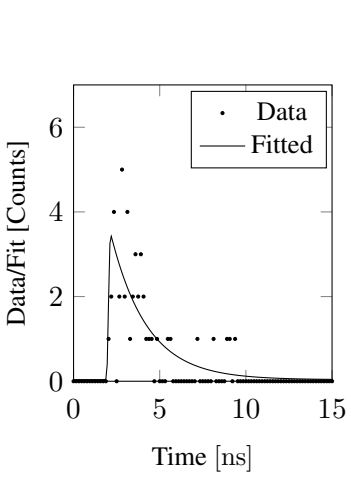
Bayesian mono-exponential decay analysis

In this chapter the results of the application of the mono-exponential Bayesian analysis developed in Part II are presented. In addition to recapping the main results presented in [1], some of the additional benefits that the Bayesian algorithms have over the other analysis techniques are explored in greater detail than was possible in [1]. In Section 6.1 the key features of the developed Bayesian analysis are discussed, before the performance of the mono-exponential Bayesian algorithm is compared with those of established analysis techniques with a focus on the analysis of low count data; the application of Bayesian analysis to discriminate between decay lifetimes is discussed in Section 6.1.1, the resilience of Bayesian lifetime estimates in the presence of significant background is discussed in Section 6.1.2. The application of the mono-exponential Bayesian algorithm to experimental biological data is discussed in Section 6.2.

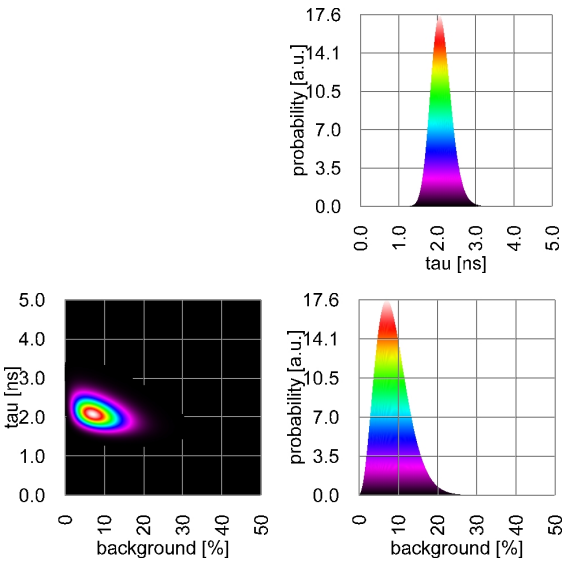
6.1 Low count synthetic data

The developed Bayesian analysis algorithms offer not only the provision of point estimates that indicate the most probable decay parameter values, in the case of Bayesian mono-exponential analysis the decay lifetime and background proportion, but the full posterior distribution, as shown in Fig. 6.1 for a synthetic data set having a total of 39 photon counts with a background proportion of 10% and a decay lifetime of 2.0 ns. The

fitted decay curve shown in Fig. 6.1 is based on the Bayesian estimated mono-exponential most-probable lifetime value of 2.17 ns and background proportion of about 8%, it being evident on inspection of the posterior distribution that there is quite significant uncertainty in the decay parameter estimates (as would reasonably be expected for data having a total of only 39 photon counts); the posterior distribution extending between lifetimes of about 1.5 ns and 3.0 ns and between background proportions of 0% and beyond 20%. The uncertainty in the estimated lifetime determined to be 0.44 ns by integrating over the posterior distribution.



(a) Low count data



(b) Posterior distribution

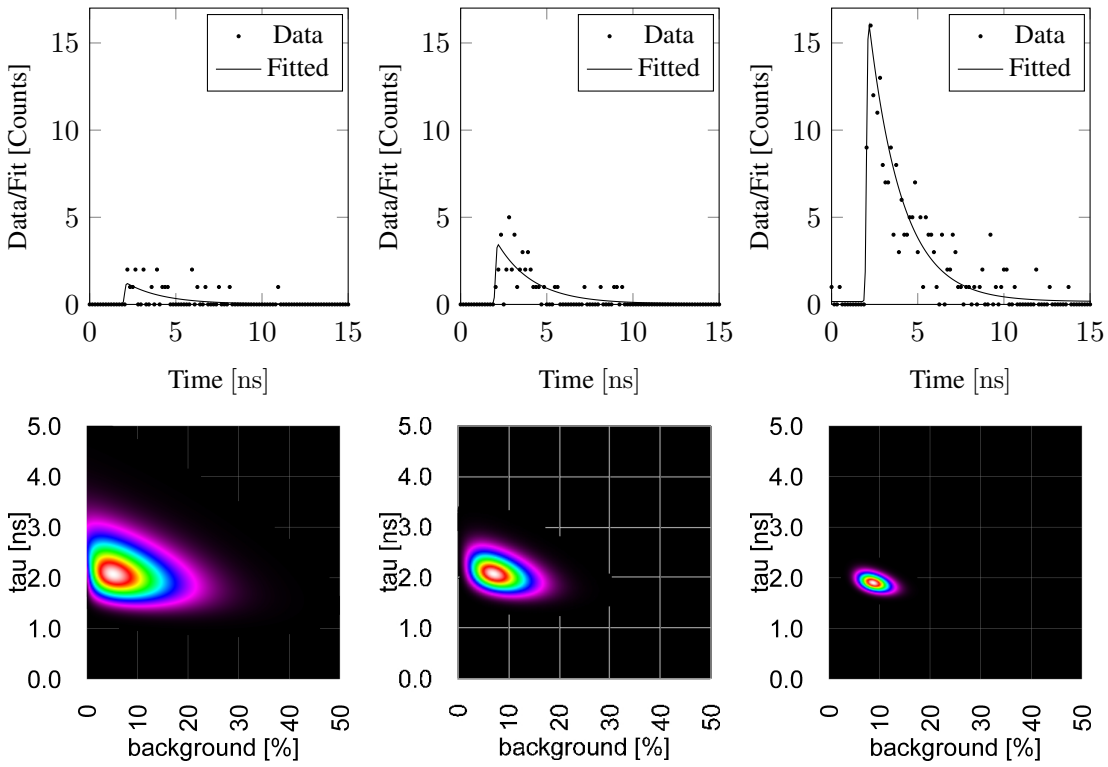
Figure 6.1: *Bayesian fitting and the posterior distribution:* In (a) simulated data having 39 total photon counts and comprising a mono-exponential decay of lifetime 2.0 ns and background count of 10% convolved with a typical instrument response, and the Bayesian predicted fitted decay using the most-probable mono-exponential decay parameter values, and in (b) the corresponding posterior distribution and its projections for the estimated background percentage and decay lifetime obtained using the Bayesian mono-exponential analysis algorithm.

The primary use of the posterior distribution is the reliable quantification of the uncertainty associated with parameter estimates, determined in this work by computation of the standard deviation of its marginal distributions. However, the posterior can also be a useful graphical aid to the analysis of decay data, especially at low total photon counts. The

potential usefulness of the posterior distribution as an easily interpreted visual guide to uncertainty in and any correlations between the estimated parameter values, and ultimately how much confidence can be placed in a set of parameter estimates, is demonstrated further in Fig. 6.2; for the analysis of synthetic data having approximately 10% background and a mono-exponential decay of lifetime 2.0 ns (convolved with a typical instrument response) the posterior distribution can be seen to concentrate around the true parameter values as the total photon count is increased for sample pixels having total counts of 39, 108, and 465 photon counts respectively. The uncertainty in the estimated lifetime, obtained by computation of the standard deviation of the lifetime marginal distribution, being determined to be 0.44 ns, 0.26 ns, and 0.12 ns for the 39, 108, and 465 total photon counts data sets respectively. It is also worthwhile noting that the posterior distribution-estimated uncertainty in both of the decay parameter estimate is in close agreement with the respective average uncertainties obtained on the analysis of many such similar data sets.

The Bayesian determined optimal single Gaussian instrument response approximation was used in performing the Bayesian pixel by pixel analysis, having been determined using the single high-count data set (about 5 million photon counts) resulting from summing the time resolved data from all of the image pixels and on the assumption of mono-exponential decay data; the optimal single Gaussian approximation having been found to have a FWHM width of 0.129 ns (i.e. a standard deviation of 0.055 ns) centered about a delay of 2.067 ns. The same single Gaussian instrument response approximation was also used in obtaining the results presented in Sections 6.1.1 and 6.1.2.

The performance of the developed mono-exponential Bayesian algorithm is compared with that of ML, LS, and phasor analysis at low total photon counts for synthetic data comprising a decay of lifetime 2.0 ns and a uniform background of 10% and incorporating an instrument response having a convolution width 0.15 ns at FWHM, as shown in Fig. 6.3; it can be observed that the Bayesian estimates converge towards the true lifetime value much more rapidly than the those of the other estimation methods. Mono-exponential Bayesian analysis yields accurate decay lifetime estimates with a precision (measured according to the standard deviation of the lifetime distribution) of about 10% at intensity of about 200 total photon counts; a precision of about 15% is offered by ML and phasor analysis at such a total photon count, and of about 20% by LS analysis. Intensities of about 400 and 450 total photon counts are needed to achieve a precision of about 10% us-



(a) $p = 39$ (b) $p = 108$ (c) $p = 465$

Figure 6.2: *The posterior as a guide to parameter estimate uncertainty:* Mono-exponential Bayesian analysis of synthetic data having approximately 10% background and a mono-exponential decay of lifetime 2.0 ns (convolved with a typical instrument response). In (a) a data set having a total photon count of only 39 photon count and the fitted fluorescence decay from the Bayesian mono-exponential predicted most probable background proportion and lifetime parameter values, with the posterior distribution beneath. In (b) and (c) the same for data sets containing 108 and 465 total photon counts respectively.

ing ML and phasor analysis respectively, and more that 500 total counts using LS. Under such conditions, to achieve lifetime estimates to a precision of about 10%, the Bayesian analysis would enable a decrease in the imaging duration of about a factor of two over that required for ML or phasor analyses. The sensitivity of the lifetime estimation techniques is now compared with reference to the F value [57]. The Bayesian-determined lifetime estimates presented in Fig. 6.3 (a), offer an F value of about 1.43, and about 1.90, 2.63 and 2.24 for the ML, LS and phasor analysis estimates respectively. The Bayesian analysis lifetime estimates are more narrowly distributed than those from the other estimation

techniques (Fig. 6.3 (b)), and this is reflected in a smaller F number which indicates the greater efficiency of the technique. The potential benefit of the improved sensitivity offered by Bayesian analysis is demonstrated in Section 6.1.1, for a simulated lifetime based sorting task.

The greater precision offered by the Bayesian lifetime estimates over those of ML, LS, and phasor analysis, is demonstrated again in the next section images when applied to the task of the identification of regions in an image based on different decay lifetimes in those regions. In Section 6.1.2 the influence of background in the data on the performance of the Bayesian algorithm is discussed.

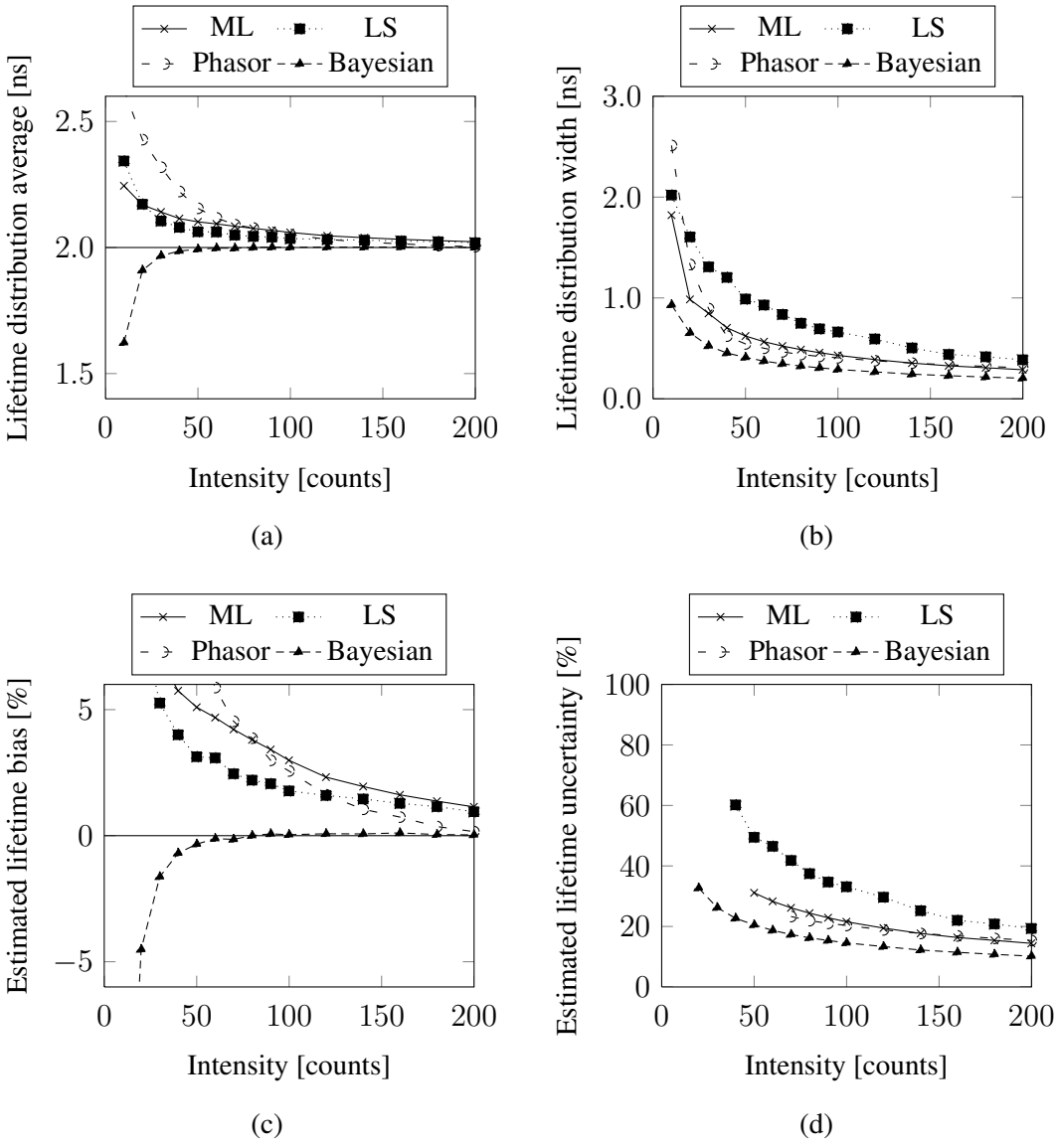
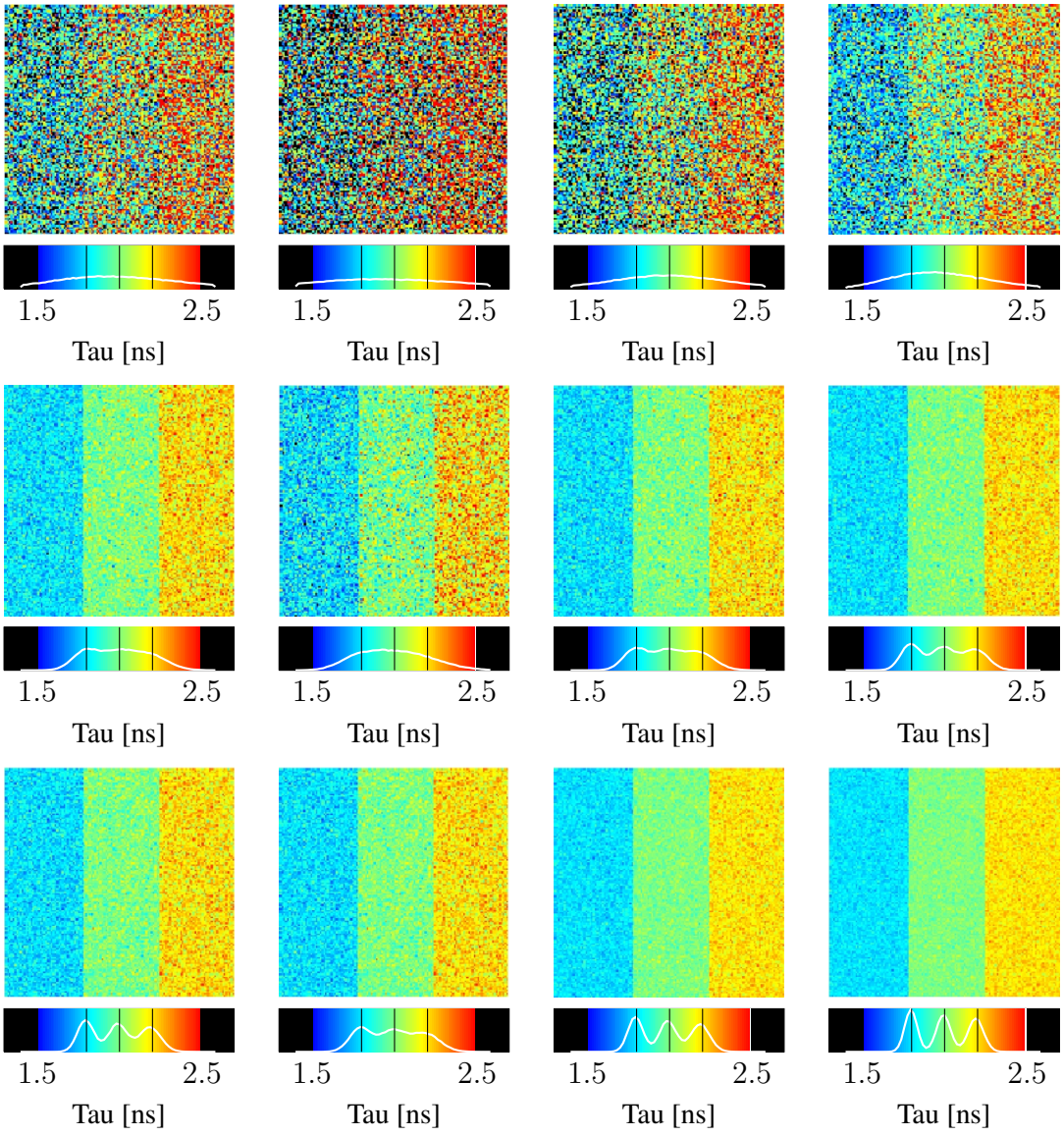


Figure 6.3: *Increasing average total photon count:* The estimated lifetime distribution statistics for increasing total count, for the different analysis methods applied to synthetic data having a decay of lifetime 2.0 ns, generated to yield an average total count of between 10 and 200 counts at each pixel with a background of 10% (convolved with a typical instrument response). In (a) the average value of the estimated lifetime distribution, in (b) the standard deviation of the estimated lifetime distribution, in (c) the bias of the lifetime estimates, and in (d) the percentage uncertainty of the estimates shown for lifetime estimates having a bias of less than 5%. The analysis was performed on synthetic 256×256 pixel images such that each data point represents the results from 2^{16} transients.

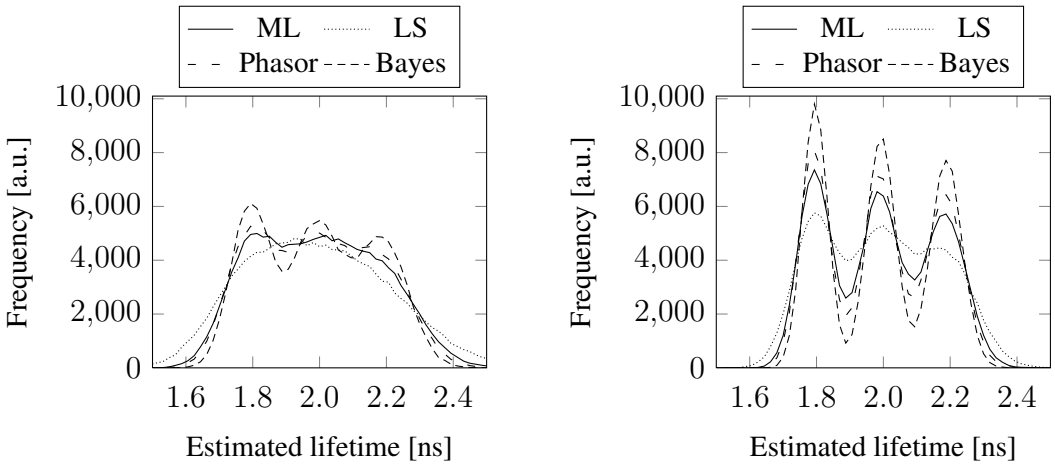
6.1.1 Lifetime-based sorting

The sensitivity of the developed mono-exponential Bayesian analysis is highlighted once more in Fig. 6.4, the lifetime estimates of images comprising three panels, with each panel separated in lifetime by 0.2 ns from its neighbour, are compared. At an average total count of $p = 100$ counts (Fig. 6.4 (top row)) the three vertical panels are very nearly discernable from each other by eye on inspecting the Bayesian analysis lifetime image, although the estimated lifetime distribution does not reflect such structure. At the higher average intensity of 1000 total photon counts (Fig. 6.4 (middle row)), the panels are clearly discernable by eye on inspection of the estimated lifetime image for each of the methods, and additionally the Bayesian analysis estimated lifetime distribution is clearly tri-modal. The contrast metric $(I_{\max} - I_{\min})/(I_{\max} + I_{\min})$ is used to quantifying the ease with which the modes of the lifetime distribution are resolved, where $I_{\max}$ is the frequency at the dominant mode and $I_{\min}$ is the lowest frequency between the modes. The contrast achieved with an average total count using our Bayesian analysis is 0.26 between the 1.8 ns and 2.0 ns panels and 0.15 between the 2.0 ns and 2.2 ns panels; the reduction in contrast between the panels with greater lifetime is due to increasing precision in the estimates with increasing measurement interval. At this total count it is not yet possible to quantify the contrast of the LS estimates as their distribution does not exhibit prominent separate modes. On increasing the average total count further it becomes possible to quantify the contrast between adjacent panels, for the sake of comparison, of all methods; at an average total count of 2500 (Fig. 6.4 (bottom row)), the contrast achieved between the 1.8 ns and 2.0 ns panels using Bayesian analysis is 0.83, a figure that compares favourably with those of ML, LS and phasor analysis estimates which yield contrast measures of 0.48, 0.18 and 0.61 respectively. It is also noticeable that the width of the estimated lifetime distributions increases with increasing lifetime; a consequence of the ratio of the decay lifetime to the measurement period [100, 74].



(a) ML (b) LS (c) Phasor (d) Bayesian

Figure 6.4: *Lifetime separation (I):* Estimated lifetime images for the analysis of synthetic images generated with decay lifetime of 1.8 ns (left panel), 2.0 ns (central panel), and 2.2 ns (right panel), for the different analysis methods, at different intensities; (from left to right) across the top row lifetime images from analysis of transients containing about 100 total photon counts for ML, LS, phasor analysis, and Bayesian analysis respectively, across the middle row the same for transients containing about 1000 total photon counts, and across the bottom row for transients having about 2500 total photon counts. The effects of an instrument response of FWHM width 0.15 ns and a uniform background averaging about 0.02 counts/bin are present in all transients.



(a) 1000 counts

(b) 2500 counts

Figure 6.5: *Lifetime separation (II):* Estimated lifetime distributions for the analysis of synthetic images generated with decay lifetime of 1.8 ns (left panel), 2.0 ns (central panel), and 2.2 ns (right panel), for the different analysis methods, in (a) for an average intensity of 1000 total photon counts, and in (b) for transients containing an average of about 2500 total photon counts.

6.1.2 Low signal to noise data

In this section the performance of the Bayesian mono-exponential analysis in poor signal to background conditions is discussed. The decay lifetime estimates obtained on the analysis of synthetic data sets having about 1000 total photon counts are shown in Fig. 6.6 for the different analysis methods for a range of signal to background proportions, the extraction of a mono-exponential decay of lifetime 2.0 ns being an increasingly challenging task as the background proportion is increased. On inspecting Fig. 6.6 (a), it is evident that the Bayesian lifetime estimates remain closer to the 'true' lifetime as the background proportion is increased; even when about half of the total counts are due to background the Bayesian lifetime estimates are distributed around an average value of 2.00 ns (the 'true' value) with a standard deviation of 0.17 ns, whereas ML, LS, and phasor analysis, all systematically overestimate the decay lifetime slightly with such a background proportion. Although, of course, the precision of the mono-exponential Bayesian lifetime estimates suffers as the decay signal becomes swamped by an increasing background, as illustrated in Fig. 6.6 (b), it is clear that the mono-exponential Bayesian analysis retains an advantage over the other analysis techniques and offers more accurate and more precise lifetime estimates regardless of the background proportion; for example, the ML, LS, and phasor analysis lifetime estimates distributions have standard deviation of 0.23 ns, 0.26 ns, and 0.39 ns respectively. Notice also that the accuracy of phasor analysis lifetime estimates degrades rapidly in the presence of background counts, most likely a symptom of a background offset being determined at each pixel from the (very limited) data that precedes the rise of the transient.

The superior performance of the developed Bayesian algorithm is even more apparent at very low intensities, as illustrated in Fig. 6.7, which compares lifetime estimates obtained with synthetic data having about 50 total photon counts and increasing background, for the different analysis methods; observe that the Bayesian analysis lifetime estimates do not deviate significantly from the actual lifetime value even when counts due to background account for about half of the total counts. Again, it is evident that the advantage of the greater accuracy in lifetime estimates offered by the developed mono-exponential Bayesian analysis is not diminished in the presence of a significant background.

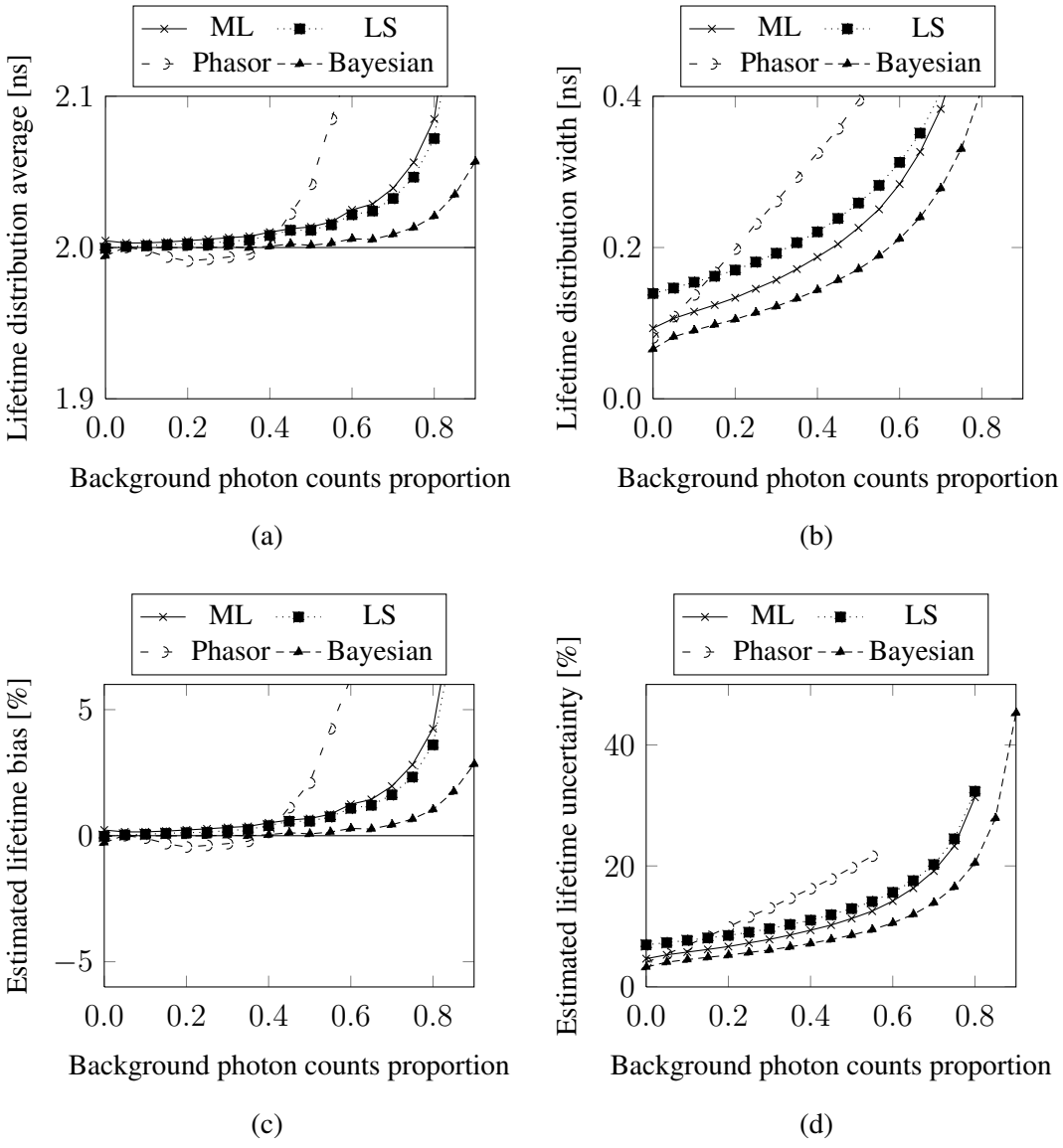


Figure 6.6: *Influence of background on lifetime estimates (Intensity: 1000 counts):* Statistics of the estimated lifetime distribution for increasing background proportion for an average total photon count of 1000 counts, for the different analysis methods. In (a) the average estimated value of the estimated lifetime distribution, in (b) the standard deviation of the estimated lifetime distribution, in (c) the bias of the lifetime estimates, and in (d) the percentage uncertainty of the estimates shown for lifetime estimates having a bias of less than 5%. In all cases the analysis was performed on synthetic 256×256 pixel images, generated to yield data corresponding to a mono-exponential decay of lifetime 2.0 ns at each pixel, and incorporating an instrument response with delay 2.0 ns and FWHM width 0.15 ns.

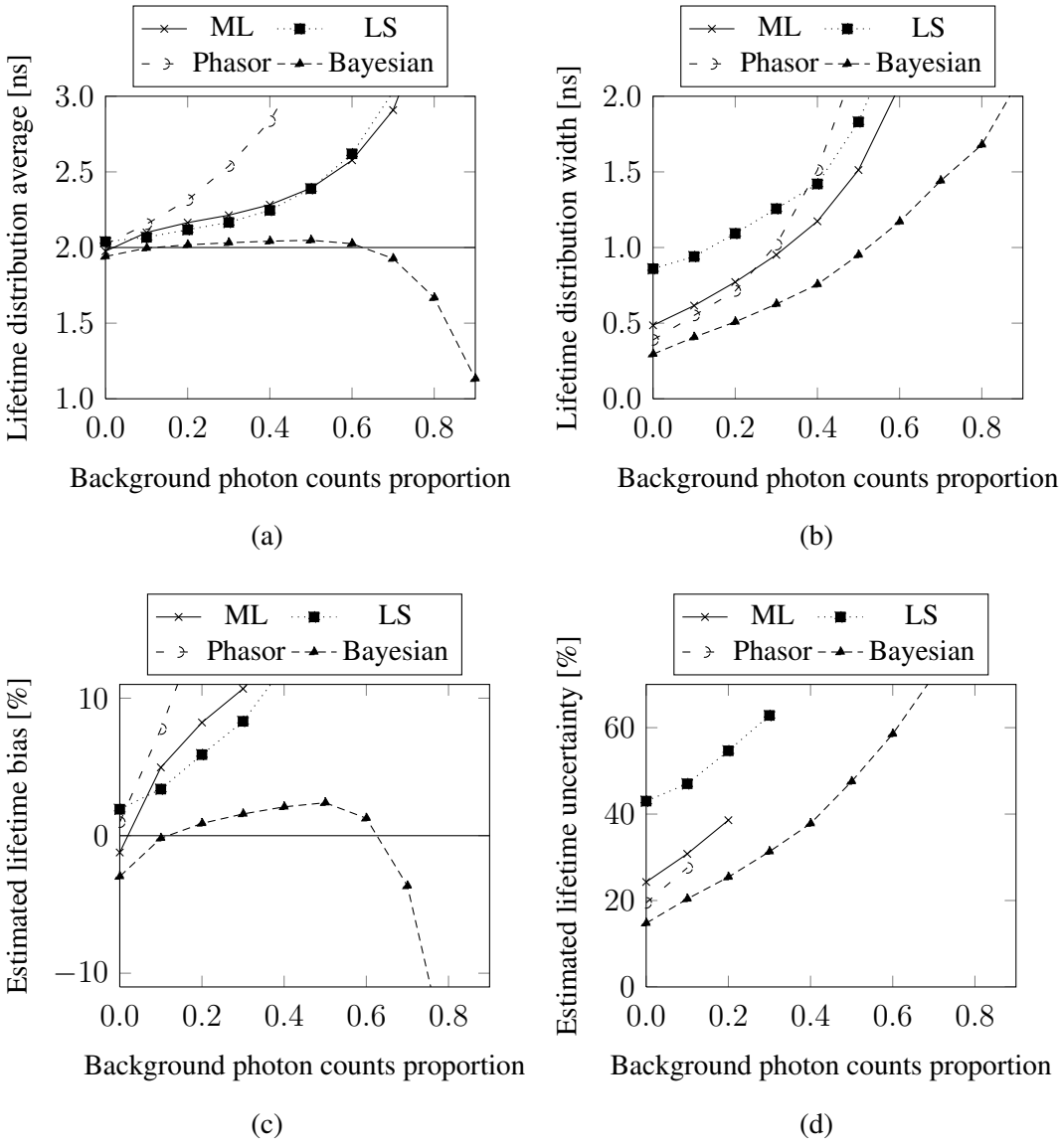


Figure 6.7: *Influence of background on lifetime estimates (Intensity: 50 counts):* Statistics of the estimated lifetime distribution for increasing background proportion for an average total photon count of 50 counts, for the different analysis methods. In (a) the average estimated value of the estimated lifetime distribution, in (b) the standard deviation of the estimated lifetime distribution, in (c) the bias of the lifetime estimates, and in (d) the percentage uncertainty of the estimates shown for lifetime estimates having a bias of less than 10%. In all cases the analysis was performed on synthetic 256×256 pixel images, generated to yield data corresponding to a lifetime $\tau = 2.0$ ns decay at each pixel, and incorporating an instrument response with delay 2.0 ns and FWHM width 0.15 ns.

6.2 Low count data from human epithelial carcinoma cells

In this section the results of the application of the mono-exponential Bayesian algorithm to experimental biological data obtained on imaging a well-established biological model system using cells which express GFP, having a fluorescence emission that follows a mono-exponential decay with a lifetime of ~ 2.2 ns, are presented. Human epithelial *carcinoma* cells expressing cdc42-GFP were imaged for short durations as described in Appendix B.2, and the performance of the mono-exponential Bayesian algorithm being quantitatively compared with that of ML, LS, and phasor analysis, on analysis of the acquired low photon count time-resolved data (at which this work is primarily targeted). All analysis methods ultimately converge to a lifetime value of around 2.18 ns as the average total photon count increases although, inevitably, some deviations from this value will be present and the effects of this biological noise are included in the lifetime distributions shown. In addition to demonstrating the improvements in lifetime estimates that mono-exponential Bayesian analysis offers over ML, LS, and phasor analysis, as originally shown in [1], the amenability of both the posterior distribution and the Bayesian parameter estimate uncertainty maps as a graphical means of assessing the level of confidence which can be placed in associated parameter estimates is illustrated.

The Bayesian determined optimal single Gaussian instrument response approximation was used in obtaining the Bayesian results presented here, having been determined using the single high-count data set (about 35 million photon counts) resulting from summing the time resolved data from all of the image pixels and on the assumption of mono-exponential decay data; the optimal single Gaussian approximation having been found to have a FWHM width of 0.204 ns (i.e. a standard deviation of 0.086 ns) centered about a delay of 2.341 ns. In Appendix B.2, the measured IRF and the Bayesian-determined approximation are compared for the Galileo time-resolved microscope (Section 1.2.1), the system used to acquire the data presented here.

The Bayesian mono-exponential parameter estimates are shown for the analysis of a low count image containing pixels having between about 35 and 350 total photon counts in Fig. 6.8. The Bayesian decay lifetime estimates are centered around an average of 2.18 ns with a standard deviation of 0.26 ns. The distribution of the associated uncertainty in the lifetime estimates (Fig. 6.8 (e)) is peaked at about 0.28 ns and, as would be expected, bears

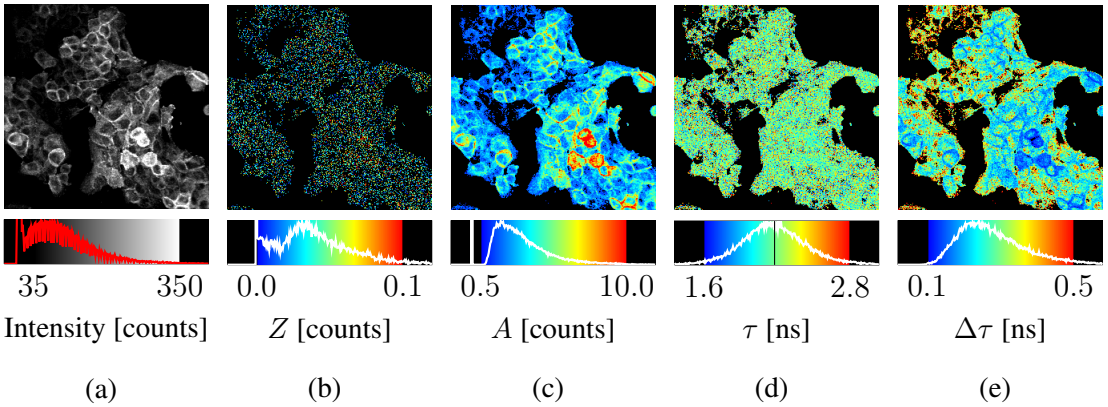


Figure 6.8: *Bayesian mono-exponential analysis of GFP data with low count data (I):* In (a) an intensity image of which pixels having a total photon count between about 35 and 350 were analysed using the mono-exponential Bayesian algorithm. In (b) and (c) the background count Z and the initial decay amplitude A estimates respectively as computed using the corresponding Bayesian estimates of the parameters w_0 and τ, and (d) and (e) the mono-exponential decay lifetime τ estimates and their estimated uncertainty $\Delta\tau$ respectively. All image panels correspond to a field of view of 334×334 μm.

a close resemblance to that of the intensity of the analysed image pixels (Fig. 6.8 (a)); regions of (relatively) high intensity having an uncertainty in the estimated lifetime as small as about 0.1 ns and pixels having low total photon counts having uncertainty in their lifetime estimates as great as about 0.5 ns. Although it is of no surprise that the precision of parameter estimates shows variation spatially within an image (it usually being the case that the more photon counts that are available for analysis the more accurate and precise are any resulting parameter estimates), Fig. 6.8 (e) does highlight that the Bayesian lifetime uncertainty map could certainly be useful as an easily interpreted graphical aid to assessing whether all, some, or even none of the lifetime estimates within an image meet the precision required for a particular investigation.

The lifetime estimates resulting from ML, LS, phasor analysis, and Bayesian analysis, are shown in Fig. 6.9 for two low total photon count images. The lifetime estimates for the very low intensity image in Fig. 6.9 (a), containing pixels that were analysed having between 25 and 175 total photon counts, are distributed (around values close to the 'true' value of 2.18 ns) with a standard deviation of 0.70 ns, 1.17 ns, 0.91 ns, and 0.41 ns for ML, LS, phasor analysis, and Bayesian analysis respectively; the tighter distribution due

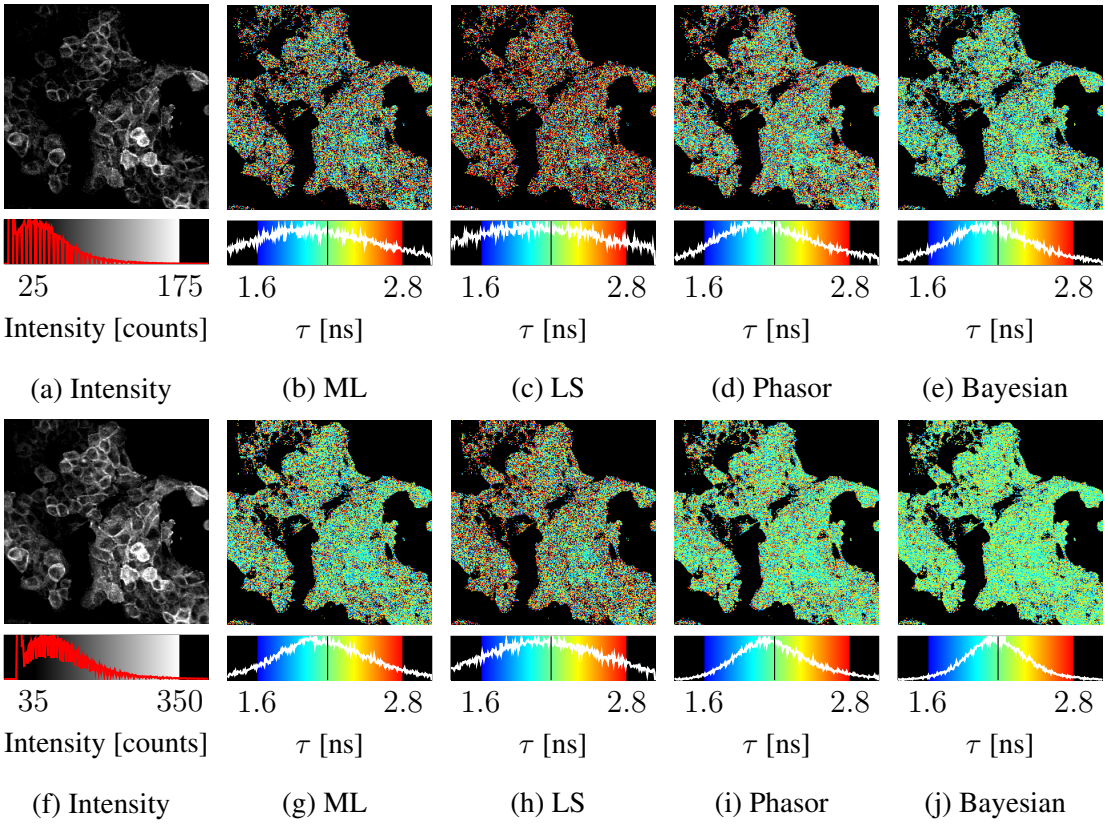
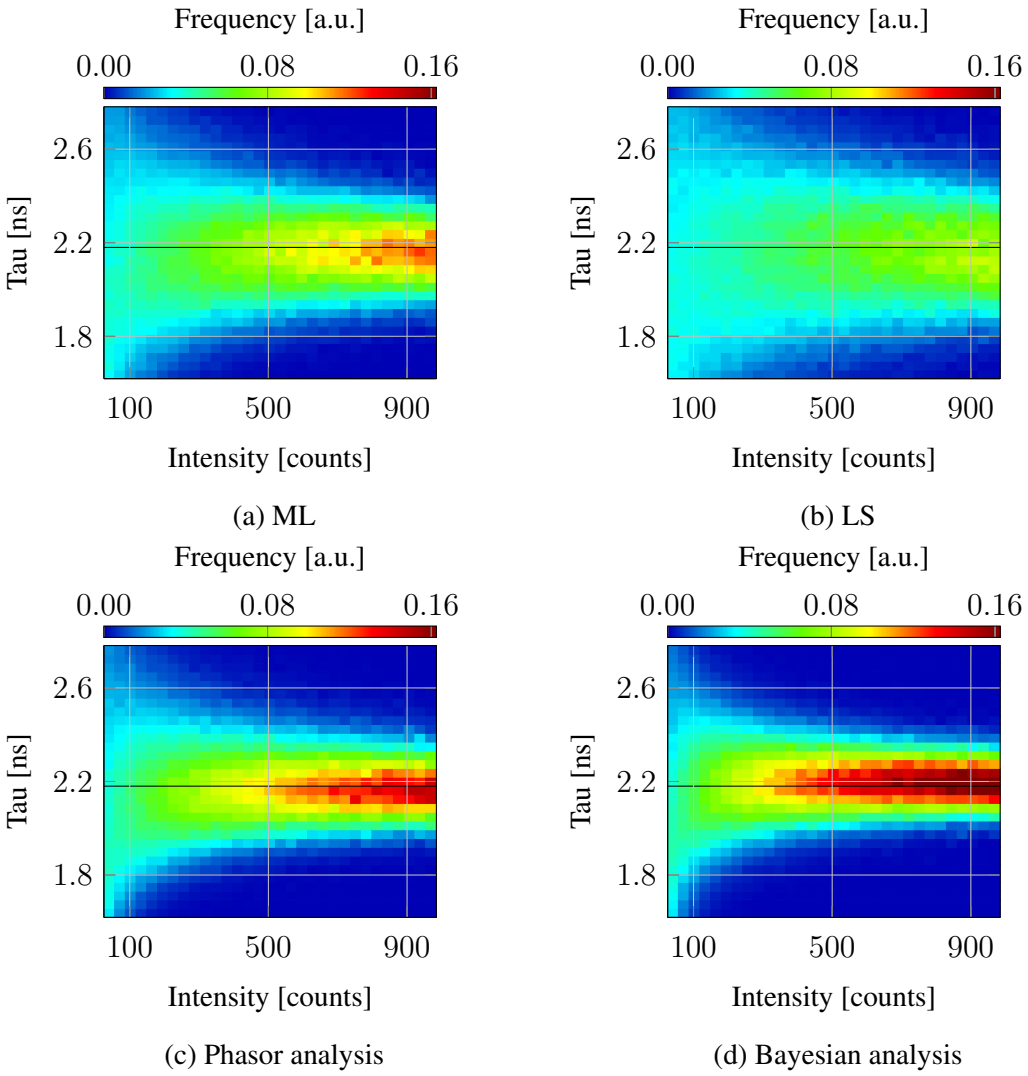


Figure 6.9: *Mono-exponential analysis of GFP data with low count data (I):* The estimated lifetime distributions for the different estimation methods applied for the analysis of GFP fluorescence decay data (2.18 ns); in (a) an intensity image of which pixels containing between about 25 and 175 total photon counts were analysed (pixels having a total count of less than 25 being masked and excluded from the analysis), and in (b), (c), (d), and (e), the mono-exponential lifetime estimates obtained by ML, LS, phasor analysis, and Bayesian analysis, of the time-resolved data. In (f) an intensity image having total photon counts between about 45 and 400, with the corresponding mono-exponential lifetime estimates due to ML, LS, phasor analysis, and Bayesian analysis, shown in (g), (h), (i), and (j) respectively. The 'true' lifetime of 2.18 ns is indicated by a black line in each lifetime image histogram. Each image panel corresponds to a field of view of 334×334 μm.

to Bayesian analysis being demonstrated by a cleaner lifetime image (Fig. 6.9 (e)) than produced using the other analysis methods (Fig. 6.9 (b,c,d)). The analysis of the intensity image in Fig. 6.9 (f) (the same data as for the Bayesian parameter estimates shown in Fig. 6.8) yield estimated lifetime distribution having width of 0.36 ns, 0.59 ns, 0.38 ns,

and 0.26 ns for ML, LS, phasor analysis, and Bayesian analysis respectively.



(a) ML

(b) LS

(c) Phasor analysis

(d) Bayesian analysis

Figure 6.10: *Mono-exponential analysis of GFP data with low count data (II):* The mono-exponential decay lifetime estimates obtained using ML, LS, phasor analysis, and Bayesian analysis, in (a), (b), (c), and (d) respectively, obtained on analysis of experimental GFP data having intensities between 10 and 1000 total photon counts. The 'true' decay lifetime of 2.18 ns, to which all methods eventually converge, is indicated by a solid line in each plot.

The mono-exponential decay lifetime estimates arising from the different analysis methods are shown in Fig. 6.10 for intensities of between 10 and 1000 total photon counts. It is evident that the Bayesian estimates are more tightly distributed around the 'true' life-

time at lower total photon counts than for the other analysis techniques; even on visual inspection the poor performance of LS analysis at low photon counts in comparison to all other techniques is striking. It is also noticeable that for this data containing negligible background phasor analysis performs more effectively than ML. As can be seen in Fig. 6.11 (a), lifetime estimates to an precision of 0.2 ns are obtained using Bayesian analysis at intensities of less than 200 photon counts, whereas ML, phasor analysis, and LS, require about 300, 400, and 800 photon counts respectively before such precision is achieved.

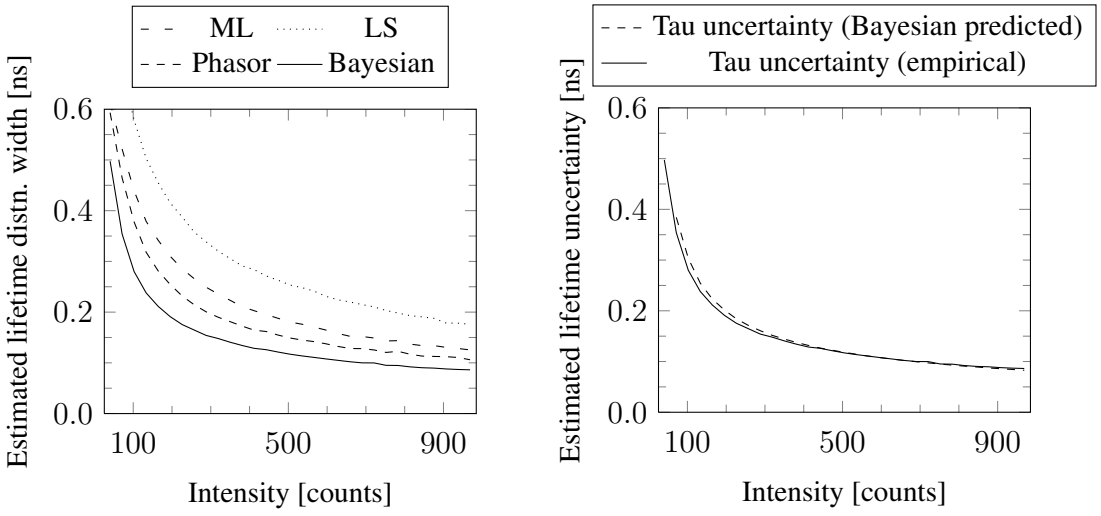


Figure 6.11: *Bayesian mono-exponential analysis for greater estimated lifetime accuracy with a reliable uncertainty measure:* In (a) the standard deviation of the estimated lifetime distributions for the different estimation methods applied for the analysis of GFP fluorescence decay data for between 10 and 1000 total photon counts; as can be seen in Fig. 6.10 Bayesian analysis offers tighter lifetime distributions than does ML, LS, and phasor analysis. In (b) a comparison of the empirically obtained uncertainty in lifetime estimates, as given by the standard deviation of the lifetime distribution obtained (i.e. the curve in (a)), and the average of the Bayesian determined lifetime error distribution.

The correspondence between the average Bayesian predicted uncertainty in a decay lifetime estimate and the empirically obtained uncertainty in the lifetime is shown in Fig. 6.11 (b); the predicted lifetime uncertainty being determined by application of the mono-exponential Bayesian algorithm to an individual data set and the empirical lifetime

uncertainty being accessible here as the lifetime across all acquired images should be uniform (barring any biological heterogeneity). It is evident that the Bayesian predicted lifetime uncertainty is in very close agreement with the lifetime estimates uncertainty determined empirically.

The potential for and the advantages of the Bayesian estimated lifetime uncertainty as an easily interpreted measure of the reliability of an analysis are highlighted for a sequence of images having increasing intensity in Fig. 6.12, this being especially relevant given that the Bayesian lifetime uncertainty prediction (obtained on analysis of individual data sets) is in such close agreement with the empirically obtained uncertainty (obtained by determining uncertainty in lifetime estimates across a large number of data sets), as shown in Fig. 6.11. It is clear, merely on visual inspection of the four Bayesian estimated lifetime uncertainty maps, that the distribution of estimated uncertainty is easily seen to shift towards lower values as the imaging duration is increased and consequently the analysed image in which the greatest confidence can be placed is very easily located. Additionally, any spatial variations that exist in the estimated lifetime uncertainty within a single image are also easily detected on inspecting the Bayesian estimated lifetime uncertainty maps; this could prove useful if only one or more regions of an image are of interest and it is only the uncertainty in the estimated lifetimes in these regions that is important. The Bayesian estimated lifetime uncertainty maps also offer the possibility of excluding image pixels that have an uncertainty in the estimated decay lifetime greater than an acceptable minimum uncertainty value from an analysis, in order that only regions of an image that meet some pre-defined level of precision can be considered.

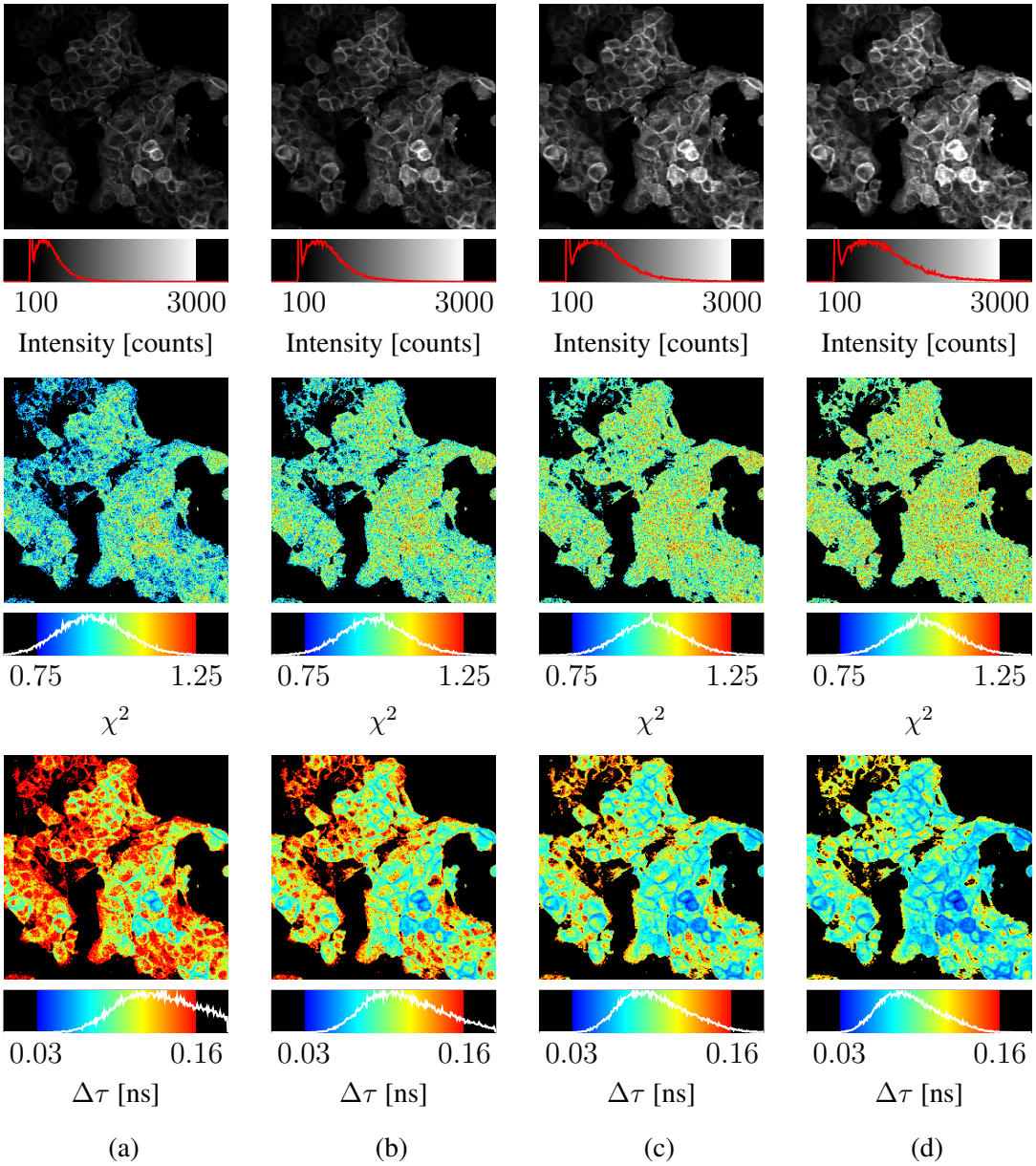


Figure 6.12: *Mono-exponential analysis of GFP data with low total count (II):* A comparison of the ML goodness-of-fit parameter χ^2 and the Bayesian estimated lifetime uncertainty $\Delta\tau$ for a sequence of images with, from left to right, increasing imaging duration. In (a) an intensity image (avg. intensity of 425 photon counts) with the corresponding ML goodness-of-fit and Bayesian estimated lifetime uncertainty maps directly below. In (b) the same for an intensity image acquired on imaging the same field for a slightly longer duration (avg. intensity of 641 photon counts), and in (c) and (d) avg. intensities of 835 and 1045 photon counts respectively.

Chapter 7

Bayesian bi-exponential decay analysis

In this chapter the performance of the Bayesian bi-exponential decay analysis algorithm is evaluated and compared with that of the ML and LS analysis methods and with global analysis. In performing bi-exponential decay analysis the objective is to accurately quantify the lifetime and the contribution to the overall decay of both decay components. In FRET studies using FLIM, biologically relevant quantities are often obtained using the bi-exponential decay parameter estimates. The results discussed in this chapter are presented with reference to the FRET efficiency and the interacting fraction, quantities that can be determined using the lifetime and initial amplitude estimates of the two decay components, as defined below for a bi-exponential decay with one decay component having lifetime τ_1 and initial amplitude A_1 and the other having lifetime τ_2 and initial amplitude A_2:

- FRET efficiency: $E = 1 - \tau_2/\tau_1, \;\; \tau_1 > \tau_2$.
- Interacting fraction: $F_2 = A_2/A_1$.

In determining quantities such as the FRET efficiency and interacting fraction using the estimated bi-exponential decay parameters values it should be expected that their accuracy and precision be limited by the accuracy and precision of the estimates on which they depend. Of course, some bi-exponential decays are more amenable to accurate analysis than others; intuitively, it would be reasonable to expect it to be more difficult to resolve the two decay components if they have similar lifetimes (i.e. a low FRET efficiency) or if one of the components dominates the decay (i.e. a very high or very low interacting

fraction). In considering bi-exponential decay analysis it is acknowledged that an exhaustive comparison of the different analysis techniques over the entire extent of the parameter space has not been undertaken, rather, the comparisons have been targeted at those regions of the parameter space that may prove most interesting. In Section 7.1 the effectiveness of the different algorithms is considered for analysis of low total count synthetic decay data for a bi-exponential decay having decay components with lifetimes that are well separated and with equal initial intensities. In section 7.1.1 the performance of the different analysis methods is compared for different FRET efficiencies, and in Section 7.1.2 the sensitivity of the different algorithms is studied as the interacting fraction is varied. The results of the application of the Bayesian bi-exponential algorithm to real experimental data is presented in Section 7.2. The data for which the results of bi-exponential parameter analysis are presented in this chapter shall also be the subject of the Bayesian decay model selection algorithm, as presented in Chapter 9.

The developed Bayesian bi-exponential decay analysis algorithm yields the most probable values for the bi-exponential decay parameters $\{w_1, w_2, \tau_1, \tau_2\}$, as defined in the time-domain FLIM system model of Chapter 4. To aid the user familiar with estimates obtained from the conventional direct fitting approach and to allow for easy comparison of the Bayesian parameter estimates with those obtained using other analysis methods, on successful completion of a Bayesian bi-exponential decay analysis, the most probable parameter values from the more familiar direct fitting model $\{Z, A_1, A_2, \tau_1, \tau_2\}$ are computed and presented (both here in this work and also to the user in the developed software) using the Bayesian parameter estimates.

In what follows, the estimated value of a bi-exponential decay parameter will frequently be compared with its known value; to avoid confusion the true value of parameter x shall be denoted by $x^\star$, for example, the known synthetic value of the lifetime τ_2 shall be denoted by $\tau_2^\star$. In discussing the bi-exponential parameter estimates the two decay components shall be addressed as the slower component $\{A_1, \tau_1\}$ and the faster component $\{A_2, \tau_2\}$; at no point is any inference as to the absolute fastness or slowness of any decay component intended.

7.1 Low count synthetic data

In this section, the performance of the bi-exponential Bayesian analysis algorithm is compared to that of ML, LS, and global analysis, for the analysis of synthetic decay data having intensities between about 1000 and 10000 total photon counts; although such intensities are an order of magnitude or so greater than those which were subject to mono-exponential analysis for Chapter 6, it should be noted that such intensities would be considered to be in the low count regime for the purposes of bi-exponential analysis. The results presented below and in Sections 7.1.1 and 7.1.2 were obtained using synthetic data that simulated a TCSPC system having a repetition rate of 40 MHz (i.e. a repetition period of 25.0 ns) and a measurement interval of 20.0 ns partitioned into 256 bins of equal width; all transients incorporated the effects of a Gaussian instrument response having a FWHM width of 0.15 ns, a uniform background of 0.5 counts per bin, and the addition of Poisson noise at each bin.

The same Bayesian-determined optimal single Gaussian instrument response approximation as was used for the analysis of synthetic data presented in Chapter 6, was also used in performing the Bayesian analysis of synthetic data presented in this chapter. The optimal single Gaussian approximation, having a FWHM width of 0.129 ns (i.e. a standard deviation of 0.055 ns) centered about a delay of 2.067 ns, was determined using a single high-count data set (about 5 million photon counts).

The analysis algorithms were tested using data simulating a bi-exponential decay having a FRET efficiency of $E^\star = 0.75$ (i.e. $\tau_1^\star = 2.0$ ns, $\tau_2^\star = 0.5$ ns) and an interacting fraction of $F_2^\star = 0.5$ (i.e. $A_1^\star = A_2^\star$). The decay parameter estimates and their statistics, as presented in Figs. 7.1, 7.2, 7.3, 7.4, 7.5, 7.6 and 7.7, were obtained on the analysis of a sequence of twenty-one synthetic 32×32 pixel images uniformly distributed between 500 and 10500 total counts per pixel, such that a total of 21504 transients were analysed. In this section the performance of the Bayesian analysis algorithm, as applied independently to every pixel of an analysed image, is also gauged against that of global analysis incorporating the assumption of decay lifetime invariance across the analysed image. Independent estimates for all of the bi-exponential parameters (i.e. Z, A_1, τ_1, A_2, τ_2) were obtained for each pixel of each of the analysed images on application of the ML, LS, and Bayesian analysis algorithms. The global ML analysis algorithm yielded a single estimate for the two decay component lifetimes (i.e. τ_1, τ_2) for each image, under the assumption of the

lifetimes being invariant over each image, and independent estimates of the background and initial amplitudes (i.e. Z, A_1, A_2) for each pixel of each of the analysed images. It should be noted that the performance of global analysis depends on the size of the image being analysed and whether the assumption of decay lifetime invariance over the image is met; this assumption is met by the synthetic data analysed in this case which presents and ideal case for the global ML analysis algorithm. It must also be stressed that the accuracy and precision of the two decay component lifetimes as estimated by the global ML analysis algorithm depend not only on the intensity at each individual pixel but also on the size of the image (i.e. the number of pixels and the total photon count of the image), and so as not to introduce confusion the accuracy and precision of the two decay component lifetimes as estimated by global ML analysis are not directly compared with those obtained using the ML, LS, and Bayesian algorithms as applied independently at each pixel. The initial amplitude estimates and the resulting interacting fraction estimates, however, are more amenable to direct comparison as they are obtained for each image pixel with all of the analysis methods, their statistics are shown in Figs. 7.5 and 7.7.

The decay lifetime estimates of the different analysis methods are shown in Fig. 7.1 for intensities between about 1000 and 10000 photon counts; the estimated lifetime distributions converge to the true lifetime values ($\tau_1^\star = 2.0$ ns, $\tau_2^\star = 0.5$ ns) for all methods as the intensity is increased, although it is evident that those due to ML and Bayesian analysis do so at lower intensities than those of LS analysis. The Bayesian estimates of both decay lifetimes are more tightly distributed than those obtained using ML; the improvement is, however, not as pronounced as it was for mono-exponential analysis. Global ML analysis produced decay lifetime estimates consistent with the true lifetimes for each of the analysed synthetic 32×32 pixel images, though demonstrated a tendency to overestimate the faster lifetime τ_2 (not shown).

The slightly greater precision, as measured by the width of the lifetime distributions (Fig. 7.1), of the Bayesian lifetime estimates is shown in Fig. 7.2, which shows the width of the distribution of lifetime estimates for intensities between about 1000 and 10000 photon counts for each of the different analysis methods; at an intensity of about 5000 total counts, the Bayesian analysis estimates of τ_1 are distributed with a standard deviation of 0.10 ns, the ML and LS estimates being distributed with standard deviation of 0.11 ns and 0.12 ns respectively. The Bayesian estimates of the lifetime τ_2 are also more tightly distributed than those obtained using ML and LS, having a standard deviation

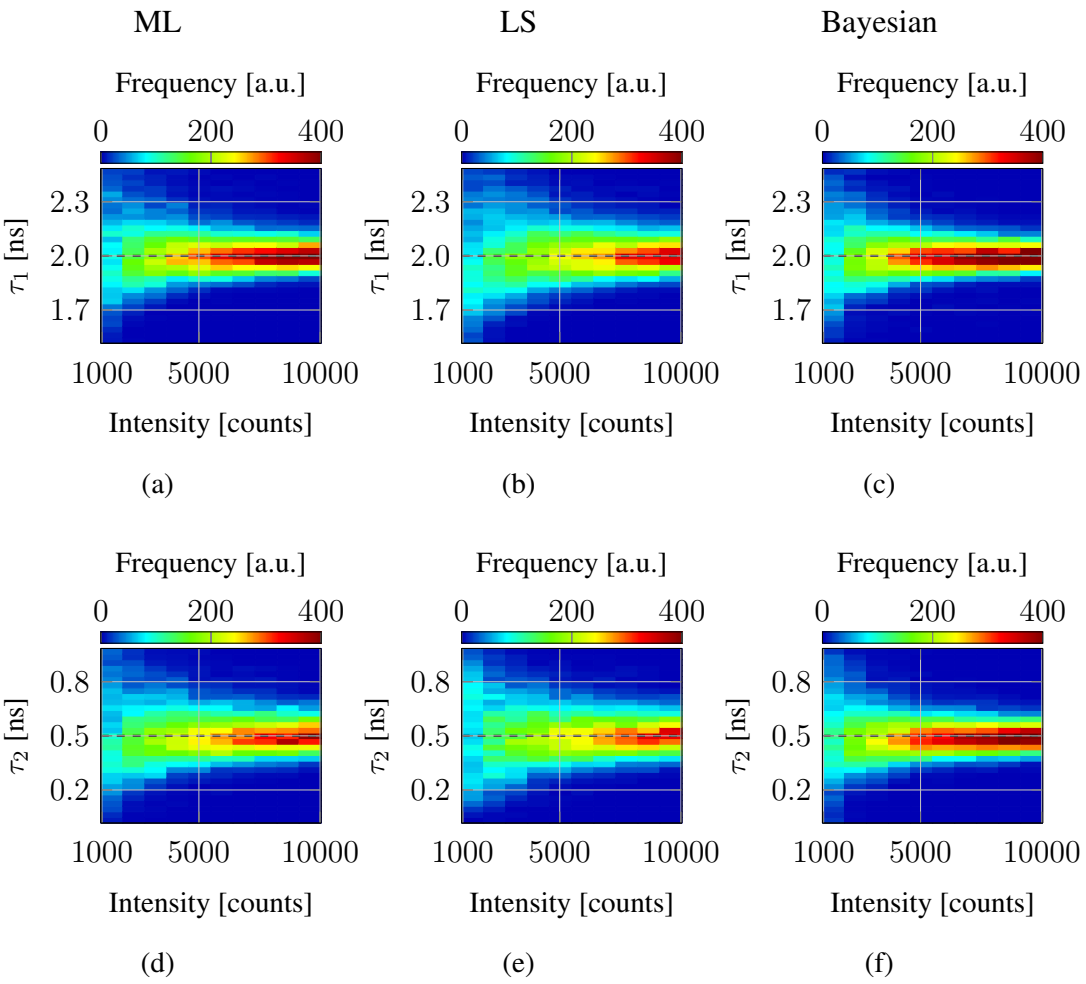
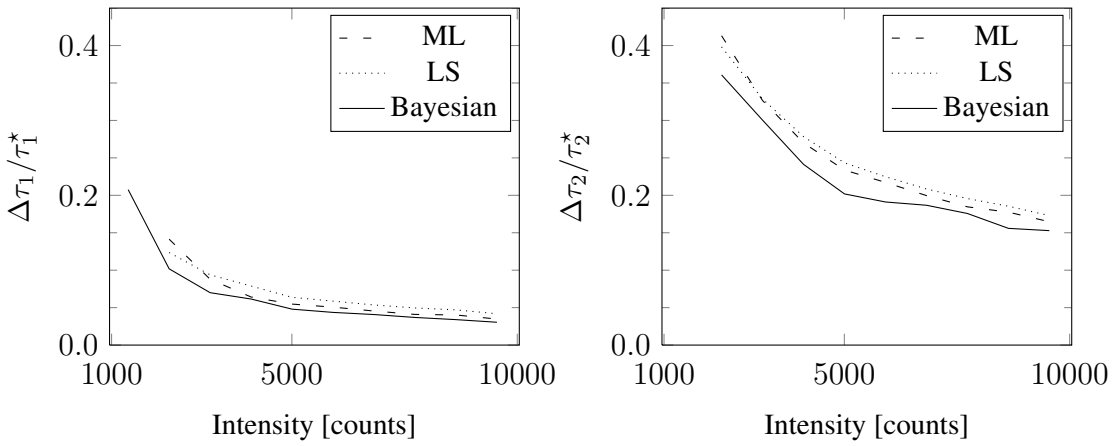


Figure 7.1: *Bi-exponential decay lifetime estimation at low counts:* Bi-exponential decay lifetime estimates obtained using ML, LS, and Bayesian analysis for the analysis of synthetic data simulating a bi-exponential decay having lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns, present with equal initial amplitudes (i.e. $A_1^\star = A_2^\star$), for intensities between about 1000 and 10000 total photon counts. In (a), (b), and (c), the estimated lifetime τ_1 obtained with ML, LS, and Bayesian analysis respectively, and in (d), (e), and (f), the estimated lifetime τ_2 obtained with ML, LS, and Bayesian analysis respectively with the true lifetime being indicated by a dashed line in each plot. Notice that ML and Bayesian analysis lifetime estimates are more precise than those of LS.

of 0.10 ns compared with about 0.12 ns for both ML and LS. Although on inspecting Figs. 7.1 and 7.2 the advantage that Bayesian analysis offers over ML and LS may appear meagre, it is worthwhile highlighting that to achieve an uncertainty of about 0.10 ns in



(a) Slower decay component

(b) Faster decay component

Figure 7.2: *Bi-exponential decay lifetime estimation precision at low counts:* The uncertainty, as measured by the standard deviation of the estimated lifetime distribution, in the bi-exponential decay lifetime estimates obtained using ML, LS, and Bayesian analysis for the analysis of synthetic data simulating a bi-exponential decay having lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns, present with equal initial amplitudes (i.e. $A_1^\star = A_2^\star$), for intensities between about 1000 and 10000 total photon counts. In (a) the width $\Delta\tau_1$ of the distribution of the estimated lifetime τ_1 obtained with ML, LS, and Bayesian analysis respectively for increasing total photon count and normalised to the known lifetime $\tau_1^\star$, and in (b) the same for the lifetime τ_2 (i.e. $\Delta\tau_2$), normalised to the known lifetime $\tau_2^\star$. The Bayesian estimates are more tightly distributed for both of the decay lifetimes. In both cases, the normalised width is displayed only when the respective lifetime estimates were not biased by more than 5% of the true value.

the estimation of the lifetime of the faster decay component an intensity exceeding about 6500 total counts for ML and about 7000 total counts for LS is required, and for the slow component such precision is achieved at intensities greater than about 6000 and 8000 total counts for ML and LS respectively; the Bayesian estimates offer this level of precision at an intensity of about 5000 total counts.

The FRET efficiency estimated using the bi-exponential lifetime estimates of the different methods is shown in Fig. 7.3 for data having between about 1000 and 10000 total photon counts. The effect of the propagation of the uncertainty in the estimated lifetimes (Fig. 7.2) on the FRET efficiency estimated using the lifetime estimates of the different methods is evident; the estimates of all methods converge to the true FRET efficiency value of $E^\star = 0.75$ as the intensity is increased, though it is clear that the FRET efficiency

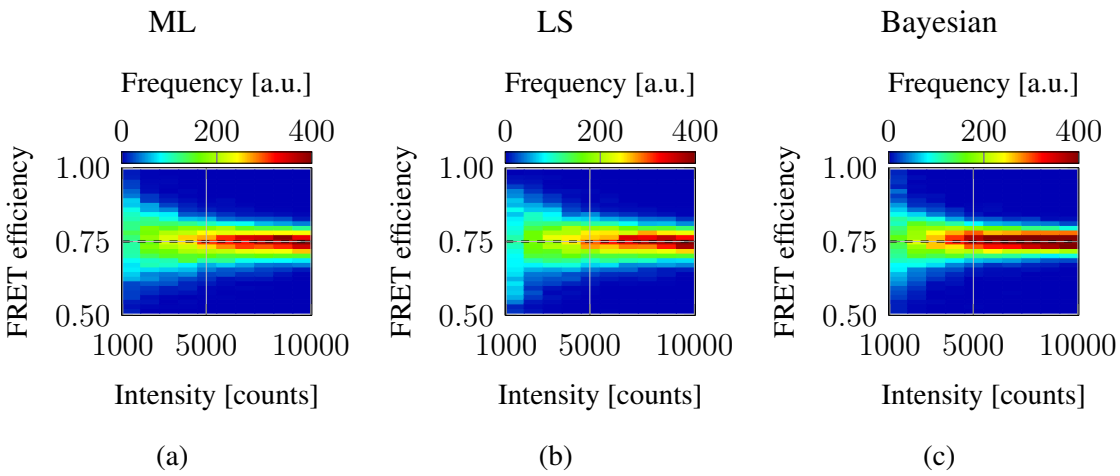


Figure 7.3: *FRET efficiency determination at low counts:* The FRET efficiency $(1 - \tau_2/\tau_1)$ as computed using the bi-exponential decay lifetime estimates obtained using ML, LS, and Bayesian analysis for the analysis of synthetic data simulating a bi-exponential decay having lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns, present with equal initial amplitudes (i.e. $A_1^\star = A_2^\star$), for intensities between about 1000 and 10000 total photon counts. In (a), (b), and (c), the FRET efficiency as determined by ML, LS, and Bayesian analysis respectively. The true FRET efficiency $E^\star = 0.75$ is indicated by a dashed line in each plot. Observe that the Bayesian estimates are more tightly distributed than those of ML and LS, a consequence of the superior decay lifetime estimates.

estimates determined using the Bayesian lifetime estimates are more tightly distributed than those of ML and LS, this being especially evident at an intensity of about 5000 total counts. The greater precision of the FRET efficiency computed using the Bayesian lifetime estimates is also demonstrated in Fig. 7.7 (a); at an intensity of about 5000 total counts the FRET efficiency derived from Bayesian lifetime estimates is distributed with a standard deviation of 0.035, whereas those derived using the estimates of ML and LS analysis are both distributed with a standard deviations of 0.047. The FRET efficiency determined from the global ML analysis decay lifetime estimates was consistent with the true value (to within 5%) for each of the analysed synthetic 32×32 pixel images (not shown); of course, such accuracy would not be achieved if the analysed images did meet the lifetime invariance assumption necessary for successful global analysis and contained, say, regions in which the decay lifetimes differed.

The estimated initial amplitude of the bi-exponential decay components are shown in Fig. 7.4; although the distributions of the initial amplitude estimates can be observed to

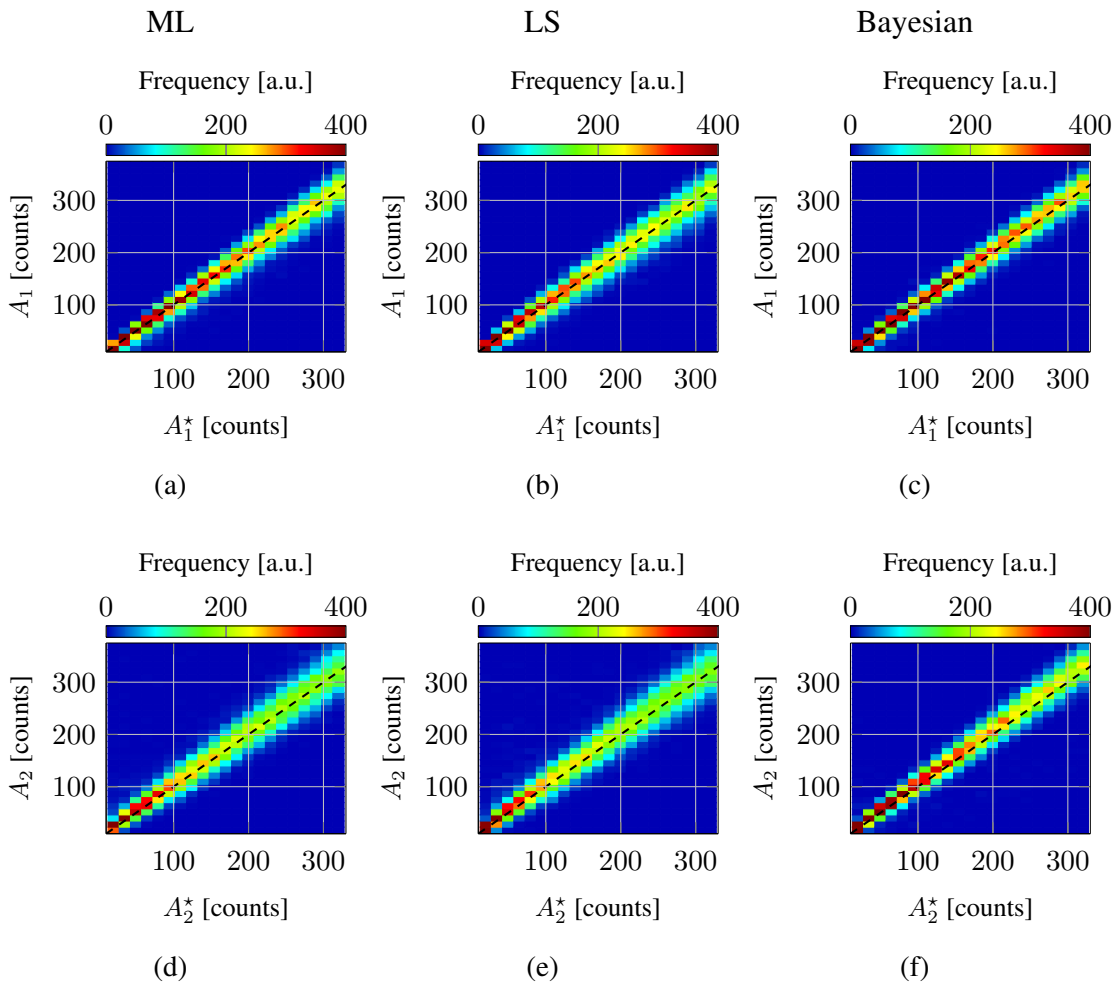
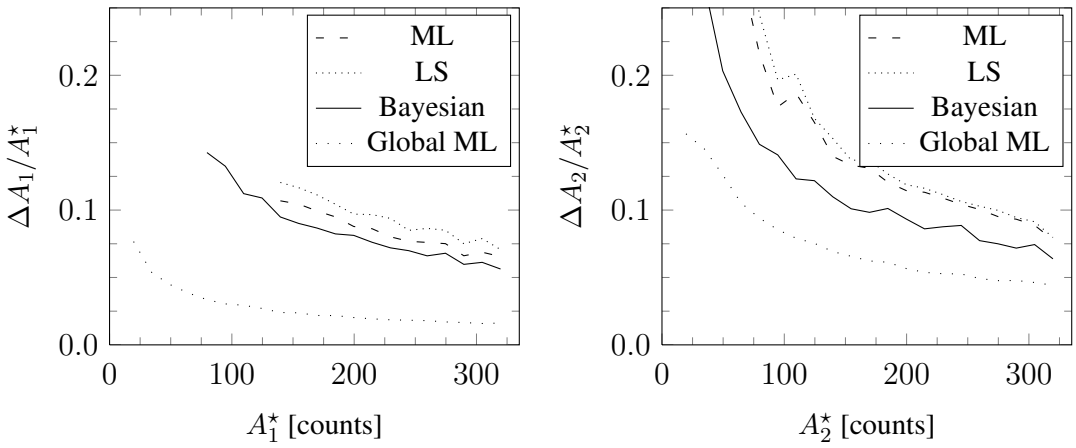


Figure 7.4: *Bi-exponential decay component initial amplitude estimation at low counts:* Bi-exponential decay component initial amplitude estimates obtained using ML, LS, and Bayesian analysis for the analysis of synthetic data simulating a bi-exponential decay having lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns, present with equal initial amplitudes (i.e. $A_1^\star = A_2^\star$), for intensities between about 1000 and 10000 total photon counts. In (a), (b), and (c), the estimated initial amplitude A_1 of the decay component having lifetime τ_1 obtained with ML, LS, and Bayesian analysis respectively, and in (d), (e), and (f), the estimated initial amplitude A_2 for the faster decay component having lifetime τ_2 obtained with ML, LS, and Bayesian analysis respectively. The true initial decay amplitudes $A_1^\star$ and $A_2^\star$ are indicated in each plot by a dashed line.

broaden as the intensity is increased, in fact the fractional uncertainty in the estimates decreases, as is demonstrated in Fig. 7.5. There is little difference between the distribu-

tions of the initial amplitude of the slower decay component for ML, LS, and Bayesian analysis methods; the estimates of the initial amplitude of the faster component due to Bayesian analysis can, however, be seen to be more tightly distributed than those due to ML and LS analysis. The precision of the global ML analysis initial amplitude estimates is superior to that offered by the ML, LS, and Bayesian analysis algorithms, all of which also independently estimates the decay lifetimes at each image pixel; the more precise initial amplitude estimates that global ML analysis offers yield superior estimation of the interacting fraction, as shown in Fig. 7.7.



(a) Slower decay component

(b) Faster decay component

Figure 7.5: *Bi-exponential decay component initial amplitude estimation precision at low counts:* The uncertainty, as measured by the standard deviation of the estimated initial amplitude distribution, in the bi-exponential initial amplitude estimates obtained using ML, LS, and Bayesian analysis for the analysis of synthetic data simulating a bi-exponential decay having lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns, present with equal initial amplitudes (i.e. $A_1^\star = A_2^\star$), for intensities between about 1000 and 10000 total photon counts. In (a) the width $\Delta A_1/A_1^\star$ of the distribution of the estimated initial amplitude A_1 of the decay component having lifetime τ_1 as obtained with ML, LS, and Bayesian analysis respectively, and in (b) the same for the initial amplitude A_2 for the faster decay component having lifetime τ_2. In all cases, the precision is displayed only for estimates that are not biased by more than 5%. Observe that the Bayesian estimates are more tightly distributed than those of ML and LS analysis for both of the estimated initial amplitudes.

The greater precision offered by Bayesian analysis in the estimation of the initial amplitudes of the bi-exponential components, particularly the initial amplitude of the faster

decay component, is illustrated in Fig. 7.5. The fractional error, as determined from the width of the initial amplitude distribution, is not markedly different for the estimation of the slower decay component for all of the methods, although Bayesian analysis is slightly superior; the Bayesian estimates of the initial amplitude of the slower decay component are distributed with a standard deviation of about 10% of its actual value when the initial amplitude exceeds about 140 counts, the same accuracy is not obtained by ML and LS analysis until the initial amplitude exceeds about 170 counts and 200 counts respectively. The estimation of the initial amplitude of the faster decay component is noticeably more precise using Bayesian analysis as compared to ML and LS analysis; estimates are distributed with a width of less than 10% of the true value for amplitudes greater than about 200 counts, the same level of precision is not obtained with ML and LS analysis until the amplitude exceeds about 275 counts.

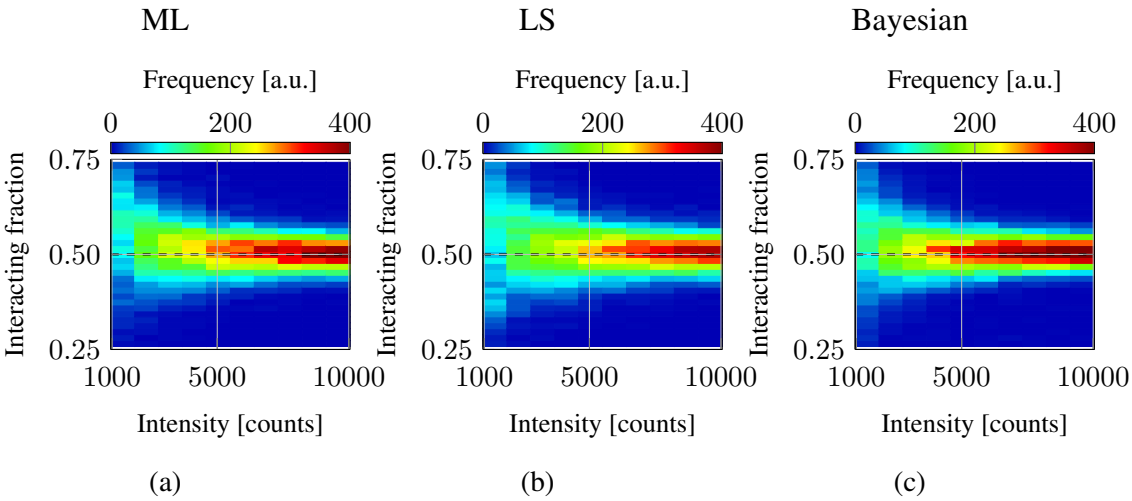
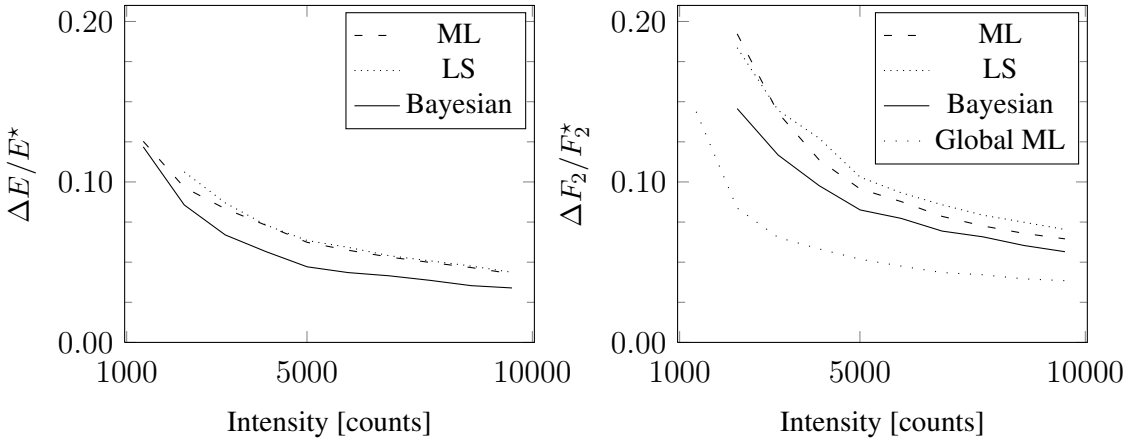


Figure 7.6: *Interacting fraction estimation at low counts:* The interacting fraction (A_2/A_1) as computed using the bi-exponential decay parameter estimates obtained using ML, LS, and Bayesian analysis for the analysis of synthetic data simulating a bi-exponential decay having lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns, present with equal initial amplitudes (i.e. $A_1^\star = A_2^\star$), for intensities between about 1000 and 10000 total photon counts. In (a), (b), and (c), the interacting fraction as determined by ML, LS, and Bayesian analysis respectively. Observe that the ML and Bayesian estimates are more tightly distributed than those of LS, a consequence of the superior estimation of the initial amplitudes A_1 and A_2. The true interacting fraction $F_2^\star = 0.5$ is indicated by a dashed line in each plot.

The interacting fraction estimates obtained using the different analysis methods are

shown in Fig. 7.6 for intensities between about 1000 and 10000 total counts. The estimates of the different methods converge to the true interacting fraction value of $F_2^\star = 0.5$ as the total photon count is increased; the Bayesian analysis and ML interacting fraction estimates being more tightly distributed around the true value than those obtained using LS analysis. The precision of the interacting fraction determined from the bi-exponential parameter estimates due to the different analysis methods is shown in Fig. 7.7 (b); an uncertainty of less than 0.05 in the determined interacting fraction is achieved with Bayesian analysis at an intensity of about 4100 counts, it is not until the intensity approaches about 5000 total counts that the same level accuracy is obtained with ML and LS analysis. The precision of the interacting fraction determined from global ML analysis initial amplitude estimates is also shown; it is clear that global ML analysis yields more precise interacting fraction estimates than do ML, LS, and Bayesian analysis. It should be noted again, though, that the synthetic data analysed for this section presents the ideal case for global analysis as the assumption of lifetime invariance is met for each analysed image; were this not the case it is likely that any global analysis algorithm would determine incorrect decay lifetimes and incorrect initial amplitudes.

In this section the performance of Bayesian bi-exponential analysis has been assessed and compared to that of ML and LS analysis and also with global ML analysis. Although the Bayesian bi-exponential analysis are superior to those obtained using the ML and LS analysis methods, the improvement over the estimates obtained using ML analysis is only marginal. The bi-exponential parameter estimates of ML and Bayesian analysis and the FRET efficiency and interacting fraction values determined from them have been demonstrated to be superior to those offered by LS. However, the marginal improvement in precision that has been demonstrated comes at the cost of a substantial increase in the time required for image analysis from seconds using ML and LS to hours using Bayesian analysis. In circumstances where the intensity of acquired data is low and the greatest possible accuracy and precision in the bi-exponential parameter estimates is required, then application of the developed bi-exponential Bayesian analysis should be seriously considered.



(a) FRET efficiency

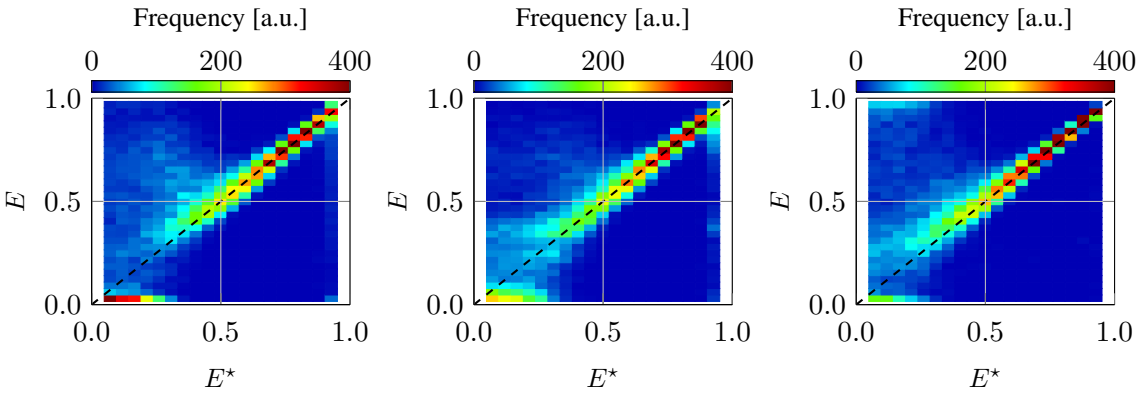
(b) Interacting fraction

Figure 7.7: *FRET efficiency and interacting fraction estimation precision at low counts:* The uncertainty, as measured by the standard deviation of the estimated parameter distributions, in the FRET efficiency $(1 - \tau_2/\tau_1)$ and the interacting fraction (A_2/A_1) estimates obtained using ML, LS, and Bayesian analysis for the analysis of data simulating a bi-exponential decay having lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns, present with equal initial amplitudes (i.e. $A_1^\star = A_2^\star$), for intensities between about 1000 and 10000 total photon counts. In (a) the width ΔE of the distribution of the FRET efficiency computed from the estimated decay lifetimes obtained with ML, LS, and Bayesian analysis for increasing total photon count, and in (b) the uncertainty ΔF_2 in the interacting fraction as estimated using the initial amplitude estimates of the different analysis methods. In all cases curves are plotted only within the range for which the estimates are biased by no more than 5%.

7.1.1 FRET efficiency estimation

In this section the performance of the different bi-exponential analysis algorithms is compared for the estimation of different FRET efficiencies, at a fixed interacting fraction. The ML, LS, and Bayesian analysis algorithms were applied to data having different FRET efficiencies, as generated by varying the ratio of the lifetimes of the bi-exponential decay components (the lifetime of the slower decay component being fixed at $\tau_1^\star$ =2.0 ns), their initial amplitudes being fixed to yield an interacting fraction of one half. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

The estimated FRET efficiency due to the bi-exponential lifetime estimates of the dif-

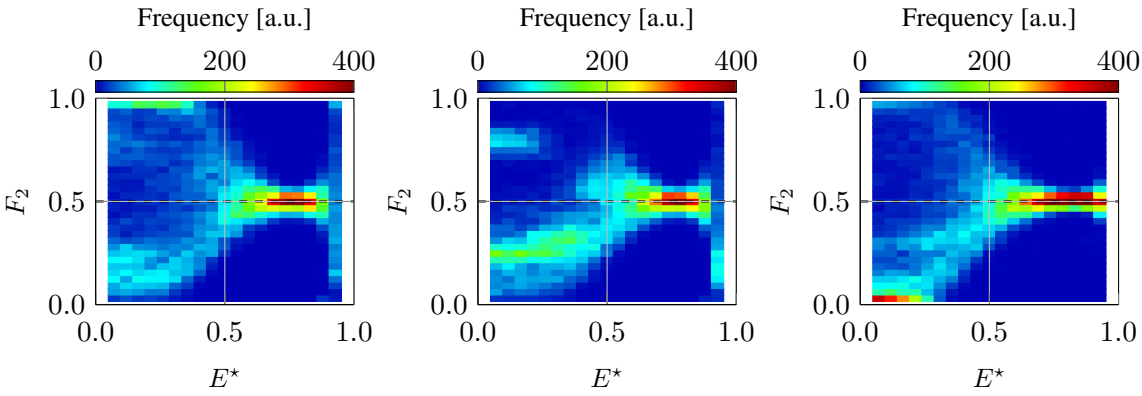


(a) ML (b) LS (c) Bayesian

Figure 7.8: *FRET efficiency estimation at different FRET efficiencies:* In (a), (b), and (c), the distribution of the FRET efficiency as determined using the bi-exponential decay parameter estimates of ML, LS, and Bayesian analysis respectively; the true value being indicated by the dashed line in each plot. Different FRET efficiencies were simulated by varying the ratio of the lifetimes of the two decay components, with equal initial amplitudes fixed to yield an interacting fraction of one half. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

ferent methods is shown in Fig. 7.8. It is evident that, for all of the different methods, the higher the actual FRET efficiency the more precisely it can be determined from the bi-exponential parameter estimates, the FRET efficiency estimates being most reliable between actual FRET efficiencies of about 50% and 90%. The distribution of the FRET efficiency estimates begins to disperse noticeably if the actual FRET efficiency is lower than about 40%, for all of the analysis methods. It is worthy of note that at a very high FRET efficiency (about 95%), when the decay lifetime of the faster component is extremely fast compared to that of the slower decay component, the distribution of FRET efficiency estimates due to Bayesian analysis is far cleaner than those due to ML and LS analysis, which both show estimates over the complete range of FRET efficiencies; seemingly ML and LS may struggle to extract reliable parameter estimates in the presence of a very fast decay component.

The interacting fraction as determined using the bi-exponential parameter estimates from the different analysis methods is shown in Fig. 7.9. It is clear that below a FRET efficiency of about 50%, at least at an intensity of about 10000 photon counts and for an actual interacting fraction of 0.5, that none of the analysis methods are able to reliably es-

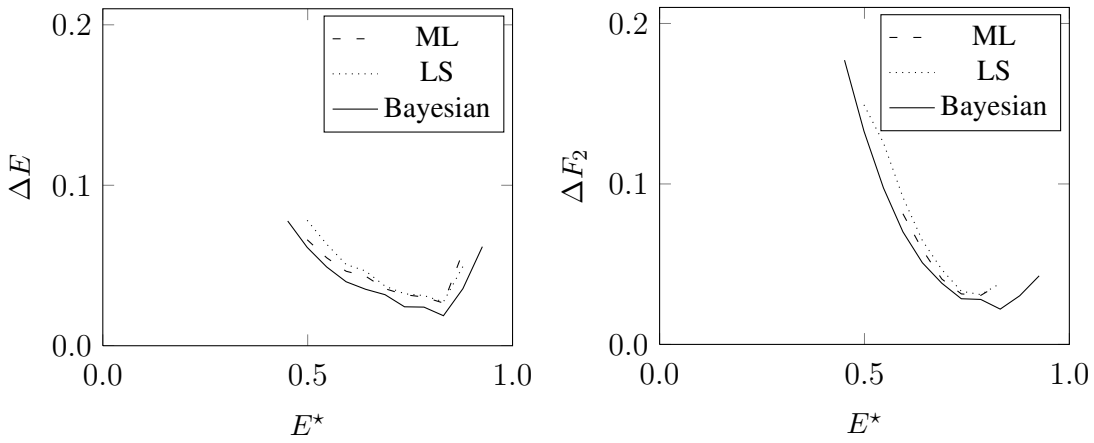


(a) ML (b) LS (c) Bayesian

Figure 7.9: *Interacting fraction estimation at different FRET efficiencies:* In (a), (b), and (c), the distribution of the interacting fraction as determined using the bi-exponential decay parameter estimates of ML, LS, and Bayesian analysis respectively; the true value of $F_2^\star = 0.5$ being indicated by the dashed line in each plot. Different FRET efficiencies were simulated by varying the ratio of the lifetimes of the two decay components, with equal initial amplitudes fixed to yield an interacting fraction of one half. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

timate the interacting fraction, with the interacting fraction being most reliably estimated between FRET efficiencies of about 50% and 90%. As was the case for FRET efficiency estimation (Fig. 7.8), the interacting fraction estimates due to ML and LS analysis are poor for the actual FRET efficiency of 95% (i.e. slower and faster decay component lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.5$ ns respectively), the corresponding Bayesian estimates, although not quite as good as those obtained at a slightly lower FRET efficiency of about 90%, remain correct.

The precision of the FRET efficiency and interacting fraction as determined using the bi-exponential parameter estimates of the different analysis methods is shown in Fig. 7.10. The uncertainty ΔE in the FRET efficiency estimates is very slightly smaller when determined using the Bayesian analysis bi-exponential parameter than when using the estimates due to ML and LS analysis, with LS analysis providing the least precise FRET efficiency estimates. The same is also true for estimation of the interacting fraction; Bayesian analysis is slightly more precise than ML, and ML is slightly more precise than LS analysis, it being clear that Bayesian analysis offers the greatest advantage for



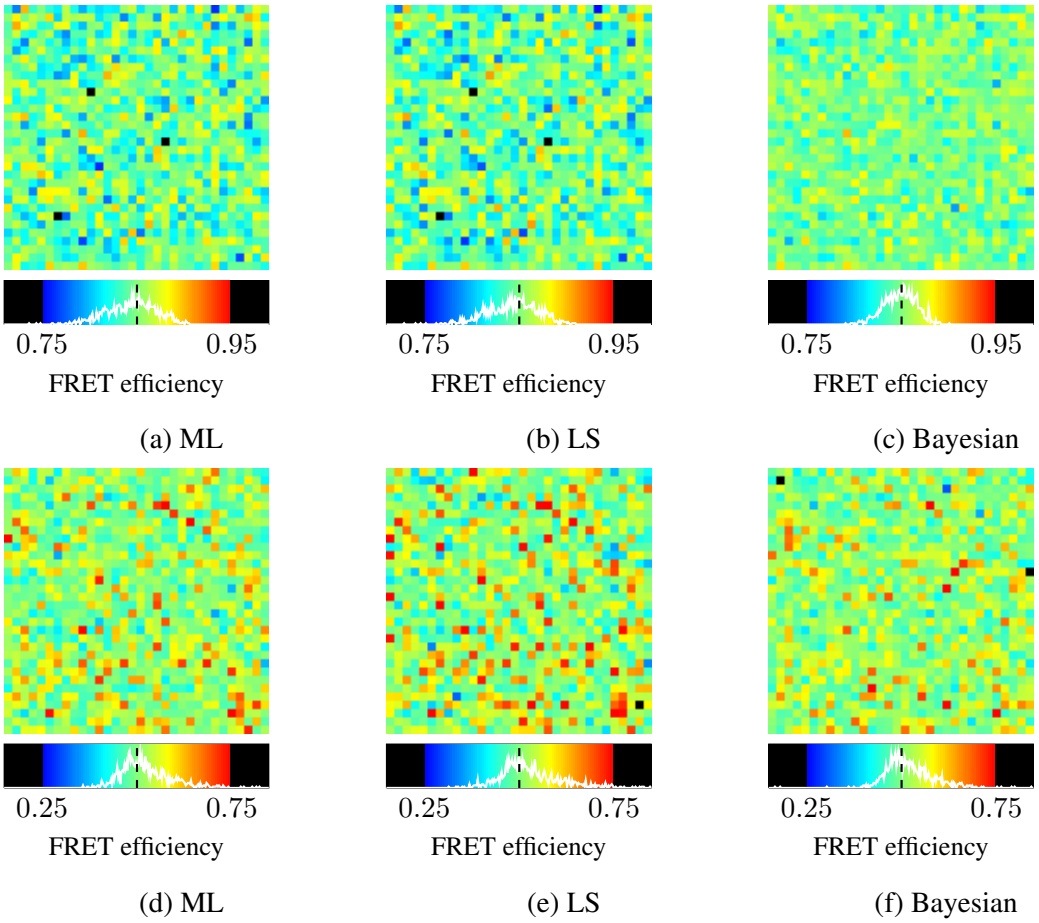
(a) FRET efficiency

(b) Interacting fraction

Figure 7.10: *FRET efficiency and interacting fraction estimation precision at different FRET efficiencies:* The uncertainty, as measured by the standard deviation of the estimated parameter distributions, in (a) the FRET efficiency $(1 - \tau_2/\tau_1)$ and in (b) the interacting fraction (A_2/A_1) estimates obtained using ML, LS, and Bayesian analysis for the analysis of data simulating a bi-exponential decay a fixed lifetime. Different FRET efficiencies were simulated by varying the lifetime of the faster decay component with the lifetime of the slower component being held fixed at 2.0 ns, the ratio of the initial amplitudes of the two decay components being equal to yield an interacting fraction of one half. All of the analysed transients were generated to have an intensity of about 10000 total photon counts. In all cases the uncertainty is displayed only for estimates having a bias that does not exceed 5%.

FRET efficiencies exceeding about 80%. The effect of the difference in precision for the different analysis methods is evidenced by the estimated FRET efficiency and interacting fraction maps at actual FRET efficiencies of 50% and 85%, as shown in Figs. 7.11 & 7.12.

The estimated FRET efficiency maps due to the different analysis methods are illustrated in Fig. 7.11 for actual FRET efficiencies of 50% and 85%. Although the difference is negligible, both of the Bayesian images are slightly less noisy than their counterparts due to ML and LS analysis. The difference being most pronounced at the higher FRET efficiency of 85%, that is when the simulated bi-exponential decay components have lifetimes $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 0.15$ ns; the FRET efficiency distribution determined using the Bayesian lifetime estimates have a standard deviation of 0.017 centered about an average value of 0.85, with the ML and LS determined FRET efficiency both being distributed with a standard deviation of 0.026 and being centered about average values of 0.85 and



(a) ML (b) LS (c) Bayesian

(d) ML (e) LS (f) Bayesian

Figure 7.11: *FRET efficiency estimation at different FRET efficiencies:* In the top row the FRET efficiency $(1 - \tau_2/\tau_1)$ maps due to ML, LS, and Bayesian analysis in (a), (b), and (c) respectively, for a FRET efficiency of 85%. In the bottom row, the same for a FRET efficiency of 50%. A FRET efficiency of 50% was simulated by setting the lifetimes of the decay components to be 2.0 ns and 0.5 ns, the ratio of the initial amplitudes of the two decay components being equal to yield an interacting fraction of one half; the faster decay component lifetime was fixed to 0.15 ns to simulate a FRET efficiency of 85%. All of the analysed transients were generated to have an intensity of about 10000 total photon counts. At each different interacting fraction 2^{10} transients were analysed.

0.84 respectively. The determination of a FRET efficiency of 50% is more challenging than that of 85% as the decay component lifetimes are closer together; using the Bayesian analysis lifetime estimates the determined FRET efficiency distribution is centered about an average of 0.52 and has a standard deviation of 0.057, the ML and LS lifetime esti-

mates yield FRET efficiency distributions having standard deviations of 0.064 and 0.073 respectively (both also being centered around an average estimated FRET efficiency of 0.52).

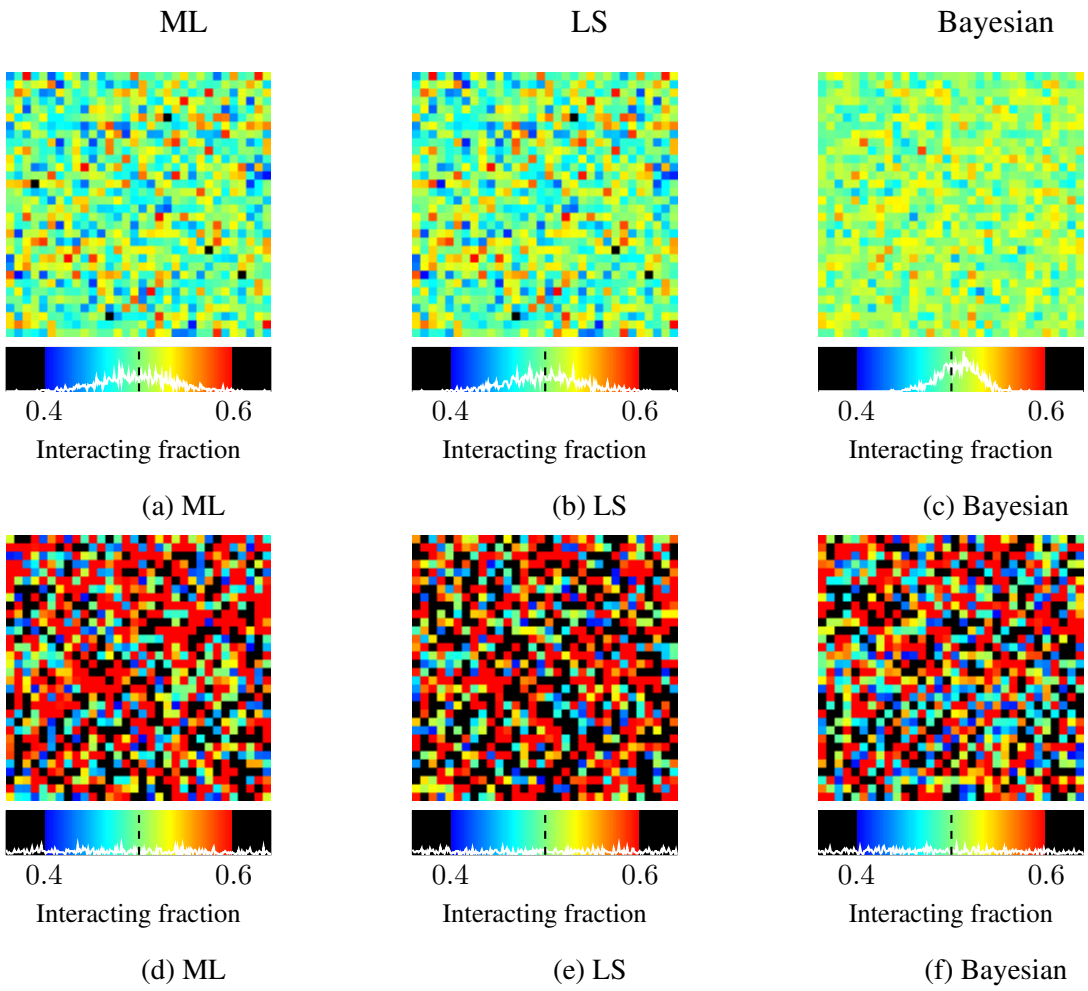


Figure 7.12: *Interacting fraction estimation and FRET efficiency:* In the top row the interacting fraction (A_2/A_1) maps due to ML, LS, and Bayesian analysis in (a), (b), and (c) respectively, for a FRET efficiency of 85%. In the bottom row, the same for a FRET efficiency of 50%. A FRET efficiency of 50% was simulated by setting the lifetimes of the decay components to be 2.0 ns and 1.0 ns, the ratio of the initial amplitudes of the two decay components being equal to yield an interacting fraction of 0.5; the faster decay component lifetime was fixed to 0.15 ns to simulate a FRET efficiency of 85%. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

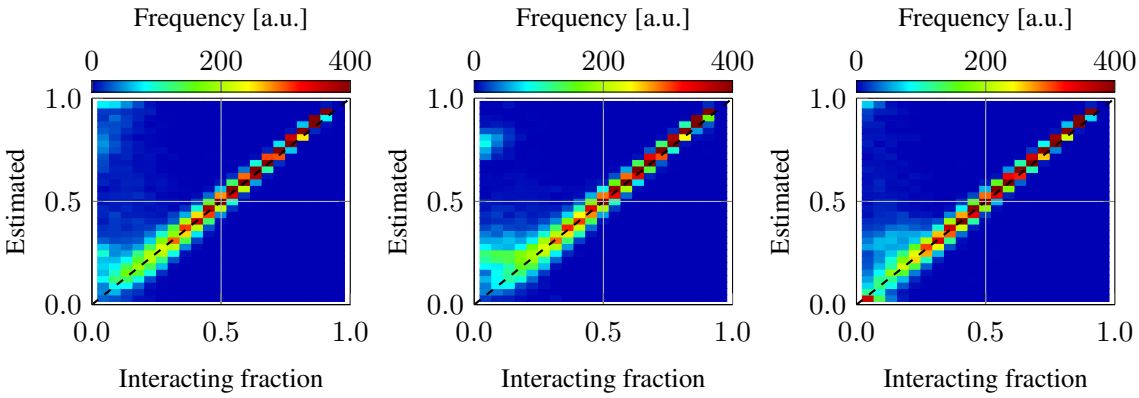
In Fig. 7.12 the interacting fraction maps due to the different analysis methods are shown for FRET efficiencies of 50% and 85%. The noticeably less noisy interacting fraction image due to the Bayesian analysis bi-exponential parameter estimates at a FRET efficiency of 85% is a consequence of the Bayesian-determined interacting fraction distribution having a standard deviation of 0.021 and being centered about an average interacting fraction value of 0.51, whereas the ML and LS distributions have a standard deviation of 0.037 and 0.038 respectively with both being centered about an average value of 0.50. The interacting fraction maps for the true FRET efficiency of 50%, that is for decay lifetimes of $\tau_1^\star = 2.0$ ns and $\tau_2^\star = 1.0$ ns, are equally poor for all analysis methods and demonstrate that obtaining reliable interacting fraction estimates at a FRET efficiency of 50% is challenging.

7.1.2 FRET interacting fraction estimation

In this section the performance of the different algorithms is compared for different interacting fractions, at a fixed FRET efficiency. The analysis algorithms were applied to data that simulated transients having different interacting fractions, generated by varying the ratio of the initial amplitudes of the bi-exponential decay components, the lifetimes of the decay components were fixed at 2.0 ns and 0.5 ns yielding a FRET efficiency of 75%. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

The distributions of the interacting fraction estimates as determined using the bi-exponential decay parameter estimates of Bayesian analysis, ML, and LS, are shown in Fig. 7.13, as the true interacting fraction is varied. The estimated interacting fraction is distributed about the true interacting fraction value with a standard deviation of less than about 0.05 for all analysis methods for interacting fractions greater than about 0.4, as shown in Fig. 7.15; the Bayesian estimates are very slightly superior to those of the other methods, and the ML estimates slightly more precise than those obtained with LS analysis. The estimation of the interacting fraction is poor when it is less than about one half, and degrades very quickly as it is decreased below about 0.3 for all analysis methods (Fig. 7.15).

The estimation of the FRET efficiency is more precise using the Bayesian lifetime estimates than those obtained using ML or LS over the tested interacting fraction range,

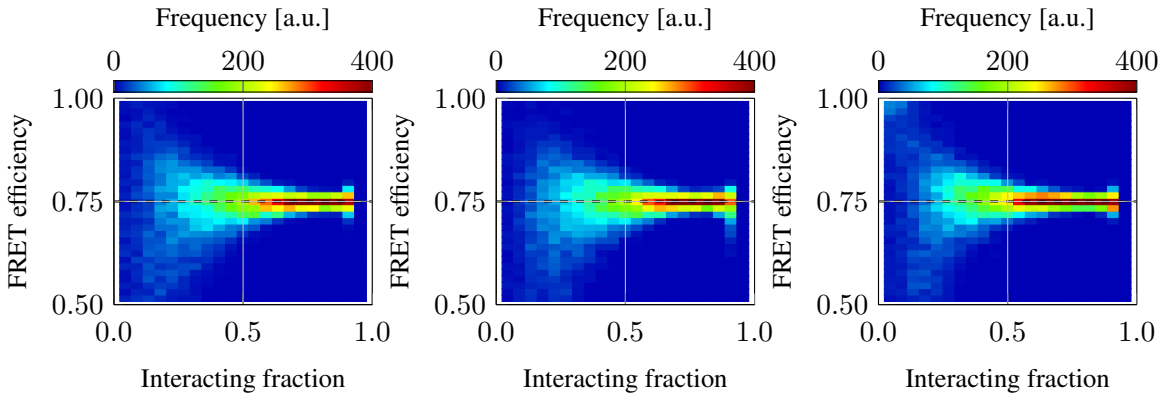


(a) ML (b) LS (c) Bayesian

Figure 7.13: *Interacting fraction estimation at different interacting fractions:* In (a), (b), and (c), the distribution of the interacting fraction as determined using the bi-exponential decay parameter estimates of ML, LS, and Bayesian analysis respectively; the true interacting fraction value being indicated by the dashed line. Different interacting fractions were simulated by varying the ratio of the initial amplitudes of the two decay components, their lifetimes being fixed at 2.0 ns and 0.5 ns. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

as can be seen in Fig. 7.14; the Bayesian determined FRET efficiency distribution remains a little tighter than those of ML and LS as the interacting fraction is decreased. The improvement in precision of the FRET efficiency is modest, however, as is also illustrated in Fig. 7.15.

There is a range of interacting fractions for which all algorithms perform effectively and yield reasonable FRET efficiency and interacting fraction estimates, as can be clearly observed on inspecting Figs. 7.13, 7.15 & 7.14. When the interacting fraction is small fewer photons from the fast decay component are counted and present in the analysed data and therefore quantification of the contribution of the fast decay component to the overall decay is more difficult than when the interacting fraction is relatively large. Similarly, when the fast decay component dominates the overall decay, that is when the interacting fraction approaches unity, relatively fewer photons from the slow decay component are counted and therefore the accuracy and precision with which its lifetime and contribution to the overall decay can be determined is degraded. It is apparent on inspecting the uncertainty in the estimated FRET efficiency, as shown in Fig. 7.15, that that the most precise estimates are obtained for interacting fractions greater than about 0.40 and less than about



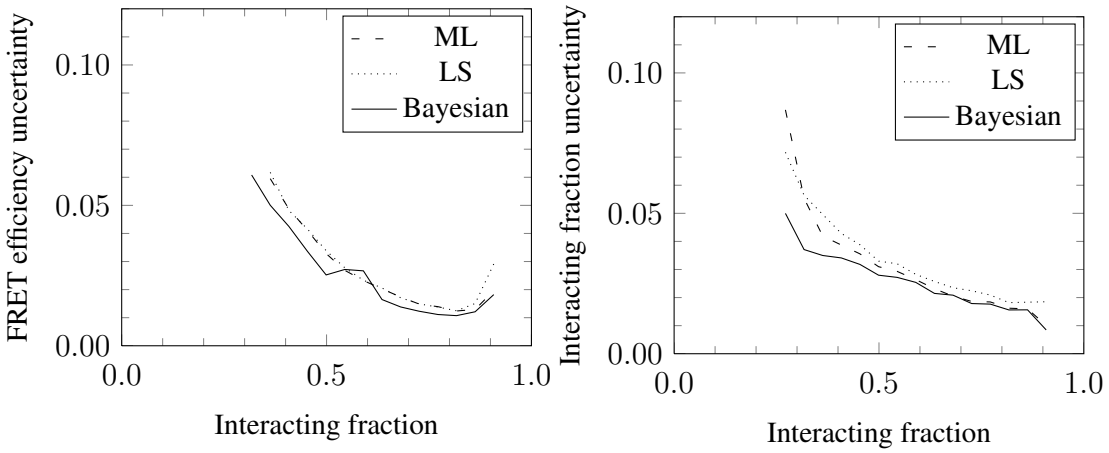
(a) ML (b) LS (c) Bayesian

Figure 7.14: *FRET efficiency estimation at different interacting fractions:* In (a), (b), and (c), the distribution of the FRET efficiency as determined using the bi-exponential decay parameter estimates of ML, LS, and Bayesian analysis respectively; the true value of 0.75 being indicated by the dashed line. Different interacting fractions were simulated by varying the ratio of the initial amplitudes of the two decay components, their lifetimes being fixed at 2.0 ns and 0.5 ns. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

0.85. The precision of the FRET efficiency estimates degrades rapidly as the interacting fraction is decreased below about 0.50 (Fig. 7.15), as can also be seen in the disperse FRET efficiency distributions shown in Fig. 7.14 for all analysis methods for interacting fractions of less than about one half. Although the accuracy and precision of both the Bayesian estimated FRET efficiency and interacting fraction estimates are consistently superior to those obtained with ML and LS over the range of interacting fractions, it is unlikely that any advantage is sufficiently significant that Bayesian analysis would be preferred over the other methods due to the time taken to perform Bayesian analysis.

7.2 Data from human breast cancer tissue

In this section the analysis of data collected as part of a study investigating intermolecular FRET in breast tissue, carried out at the Randall, King's College London, is discussed. The intensity image shown in Fig. 7.16 was acquired by imaging a slice of breast cancer tissue with directly labelled anti-ezrin IgG-Cy2 and anti-phospho PKCa (T250) IgG-Cy3 and corresponds to a 334×334 μm field of view (256×256 pixels); the Bayesian



(a) FRET efficiency

(b) Interacting fraction

Figure 7.15: *FRET efficiency and interacting fraction estimation precision:* In (a) and (b) the uncertainty, as measured by the standard deviation of the distributions shown in Figs. 7.14 & 7.13, in the FRET efficiency and the interacting fraction estimates respectively, for the different analysis methods. In all cases the uncertainty is displayed only for estimates having a bias that does not exceed 5%.

bi-exponential lifetime and initial amplitude images obtained on analysis of the corresponding time-resolved data are also shown.

In order to increase the available photon count for analysis it is sometimes necessary to apply pixel-binning when analysing an image. When performing pixel-binning, the data at each image pixel is replaced by the aggregation of the data at the pixel itself and data from neighbouring pixels. As this procedure is applied for each of the individual pixels, the total number of pixels that are to be analysed remains unchanged. To ensure sufficient photon counts for the bi-exponential analysis presented in this section, 7×7 square pixel-binning was applied, the data at each image pixel being replaced by the aggregated data from a square having a width of seven pixels and a height of seven pixels (the size of such a bin is illustrated in Fig. 7.16 (a)). Of course, in performing pixel-binning, photon counts are increased at the expense of spatial resolution as the data from a cluster of pixels is aggregated represents an average.

The bi-exponential Bayesian analysis estimates are now compared with those of ML, LS, and global ML analysis, for the analysis of the time-resolved data of the breast tissue image shown in Fig. 7.16 (a). The breast tissue sample was imaged using the Galileo

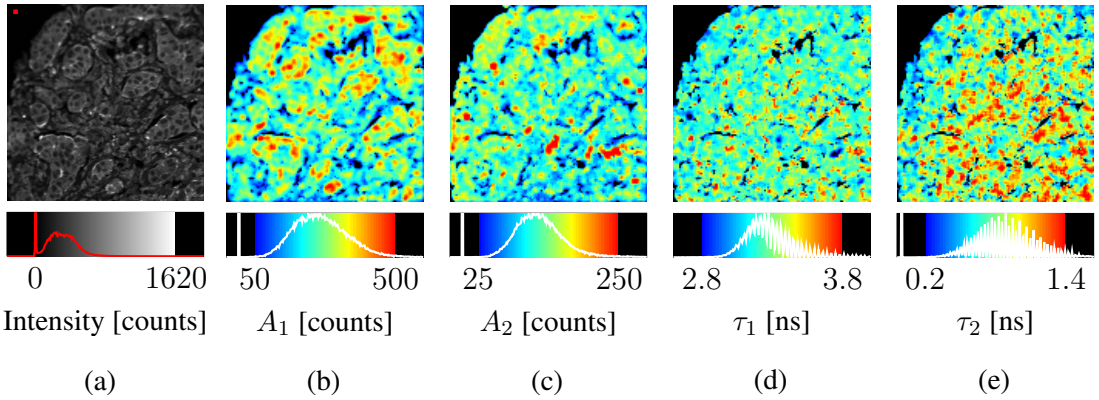


Figure 7.16: *Bayesian bi-exponential analysis of breast tissue data:* In (a) an intensity image having pixels with a total photon count of between about 130 and 1600 were analysed (having invoked 7×7 spatial binning to provide sufficient photon counts for a bi-exponential analysis) using the bi-exponential Bayesian algorithm. In (b) and (c) the estimates of the initial decay amplitudes A_1 and A_2 respectively (as computed using the corresponding Bayesian estimates of the parameters w_1, w_2, τ_1 and τ_2), and in (d) and (e) the bi-exponential decay lifetime estimates τ_1 and τ_2 respectively. The size of a 7×7 spatial bin is indicated in the top left corner of (a) by a red square. All of the images correspond to a 334×334 μm field of view, and are 256×256 pixels.

microscope, the time-resolved data being acquired into 256 time bins (of equal width) that subdivide a measurement interval of 15 ns. Image pixels having an intensity of less than 130 photon counts were masked and excluded from the analysis. As the intensity at each image pixel was insufficient for a reliable analysis to be performed, 7×7 spatial binning was invoked and the analyses were performed on time-resolved data having between about 15000 and 20000 total photon counts. The Bayesian determined optimal single Gaussian instrument response approximation was used in performing the Bayesian pixel by pixel analysis, having been determined using the single high-count data set (about 17 million photon counts) resulting from summing the time resolved data from all of the image pixels and on the assumption of bi-exponential decay data; the optimal single Gaussian approximation having been found to have a FWHM width of 0.167 ns (i.e. a standard deviation of 0.071 ns) centered about a delay of 2.255 ns.

The bi-exponential decay lifetime maps obtained on application of ML, LS, Bayesian analysis, and global ML analysis, to the time-resolved data of the imaged breast tissue sample (Fig. 7.16) are shown in Fig. 7.17. It is immediately clear that the Bayesian and

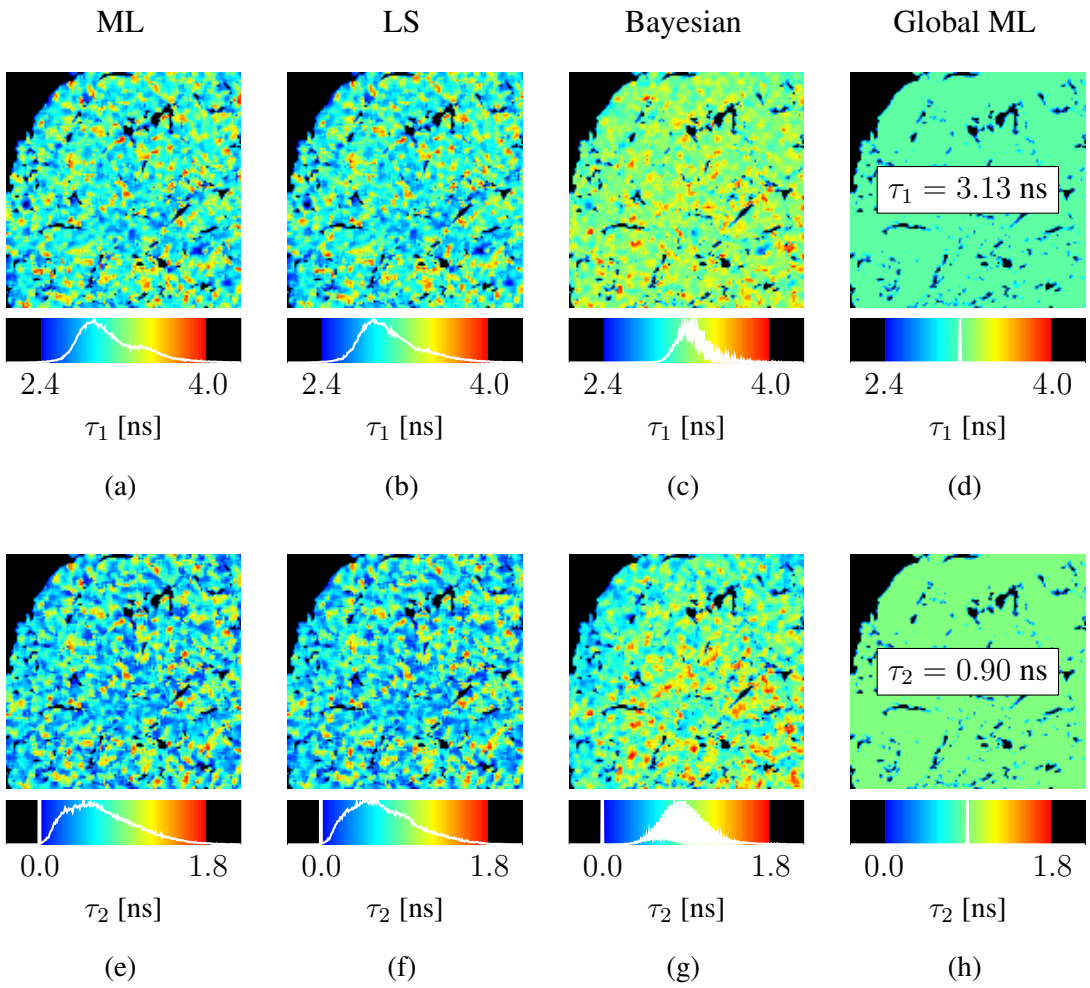


Figure 7.17: *Bi-exponential lifetime estimation:* Bi-exponential decay lifetime estimates obtained on application of ML, LS, Bayesian analysis, and global ML analysis, to time-resolved data from breast cancer tissue. In (a), (b), (c), and (d), the τ_1 lifetime maps due to ML, LS, Bayesian analysis, and ML global analysis respectively, and in (e), (f), (g), and (h) the corresponding τ_2 lifetime maps.

global ML analysis estimates of the lifetime τ_1 are considerably different to those of both ML and LS; the Bayesian estimates are, to a reasonable approximation, normally distributed around an average lifetime of 3.27 ns with a standard deviation of 0.17 ns, the distributions of ML and LS estimates being peaked at the significantly shorter lifetimes of 2.92 ns and 2.89 ns respectively despite also having a considerable number of estimates at lifetimes longer than those offered by Bayesian analysis, and global ML estimates a decay

lifetime of 3.13 ns. The distribution of the faster component Bayesian lifetime estimates is also markedly different to the distributions resulting from ML and LS analysis; the Bayesian estimates are centered around an average lifetime value of 0.76 ns, lifetimes of 0.58 ns and 0.46 ns being most frequently estimated by ML and LS respectively, and a lifetime of 0.90 ns being estimated by global ML analysis. On observing the lifetime maps resulting from the different analysis methods, it should be observed that the τ_2 maps from ML and LS are (excluding statistical noise) uniform across the image but that the Bayesian analysis τ_2 map is suggestive of regions of reduced lifetime towards the edge of the tissue core (towards the left of the image). The uniform lifetime maps due to global ML analysis demonstrate the assumption of spatially invariant lifetimes inherent to the analysis.

The bi-exponential initial amplitude estimates are shown in Fig. 7.18 for the different analysis methods. There is little difference between the A_1 estimates of the different methods, as can be seen on inspection of the initial amplitude A_1 maps. The distribution of the Bayesian initial amplitude A_1 estimates are centered around an average value of 261.4 counts, the ML and LS distributions being centered around average values of 274.7 counts and 275.5 counts respectively, and the global ML analysis estimates having an average initial amplitude of 267.5 counts. The estimates of the initial amplitude A_2 of the faster component do differ between the analysis methods; the Bayesian estimates are centered around an average initial amplitude A_2 of 123.5 counts, the ML and LS estimates having average values of 114.1 counts and and 112.4 counts respectively, and the global ML analysis estimates having an average of 91.6 counts. The Bayesian analysis estimates suggest a greater contribution from the faster component to the composite bi-exponential decay than do the estimates of the other methods.

The FRET efficiency and interacting fraction maps computed from the lifetime and initial amplitude estimates of the different analysis methods are shown in Fig. 7.19. The FRET efficiency maps due to ML and LS both suggest that the FRET efficiency is uniform across the image with no obvious regions of (relatively) low or high FRET efficiency, both methods indicating an average FRET efficiency of about 79%. The uniform FRET efficiency of 71% due to global ML analysis is a consequence of the assumption of lifetime invariance across the image. The Bayesian estimated FRET efficiency map, however, is suggestive of a higher FRET efficiency towards the edge of the tissue core than at its centre; most pixels having a FRET efficiency of about 75%. The average of the interacting

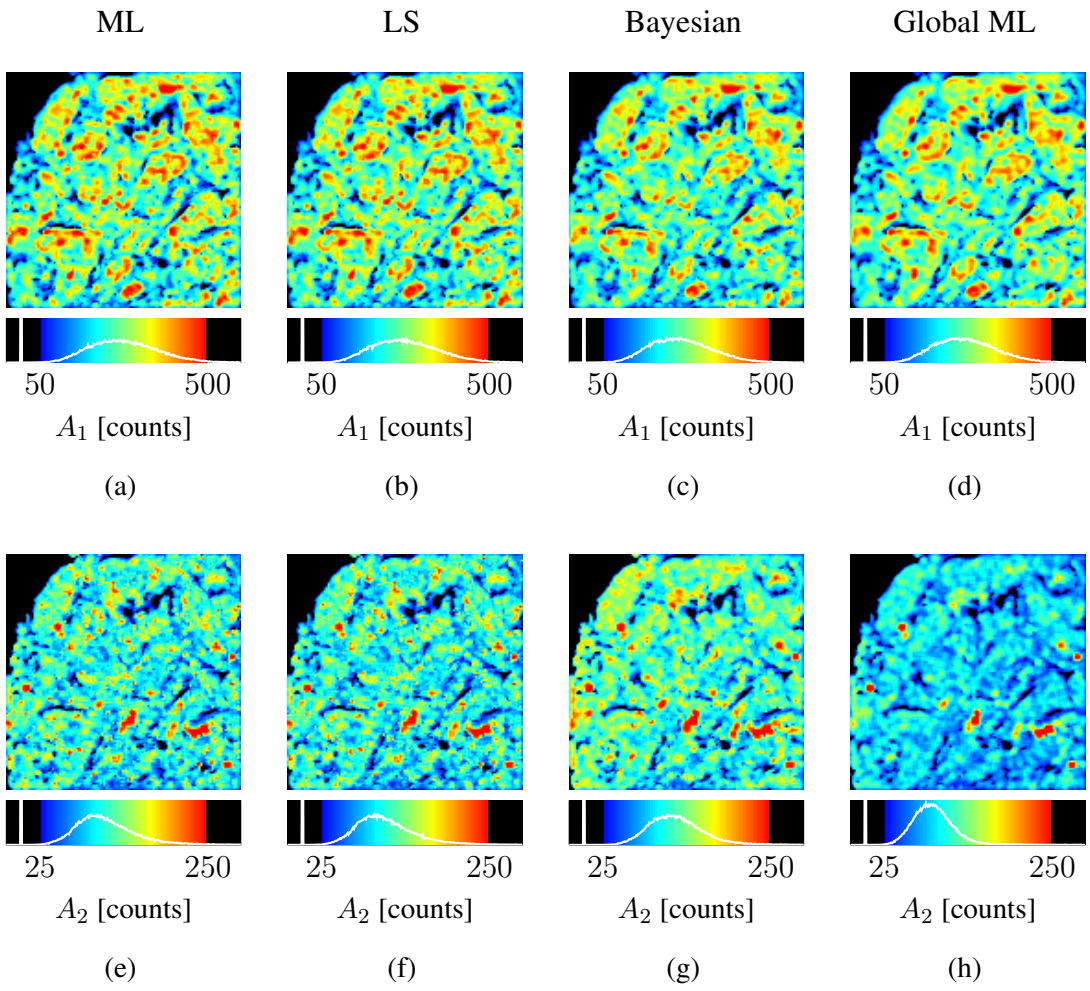


Figure 7.18: *Bi-exponential initial amplitude estimation:* Bi-exponential initial amplitude estimates obtained on application of ML, LS, Bayesian analysis, and global ML analysis, to time-resolved data from breast cancer tissue. In (a), (b), (c), and (d), the A_1 amplitude maps due to ML, LS, Bayesian analysis, and ML global analysis respectively, and in (e), (f), (g), and (h) the corresponding A_2 amplitude maps.

fraction estimates is about 29% for both ML and LS analysis, and about 32% and 26% for Bayesian analysis and ML global analysis respectively. The Bayesian, and to a slightly lesser extent the global ML analysis, interacting fraction maps suggest an increased occurrence of FRET towards the edge of the tissue core. However, neither the ML nor LS interacting fraction maps suggest that any particular regions exhibit more FRET than others. It should be noted that the band of higher FRET efficiency towards the edge of the

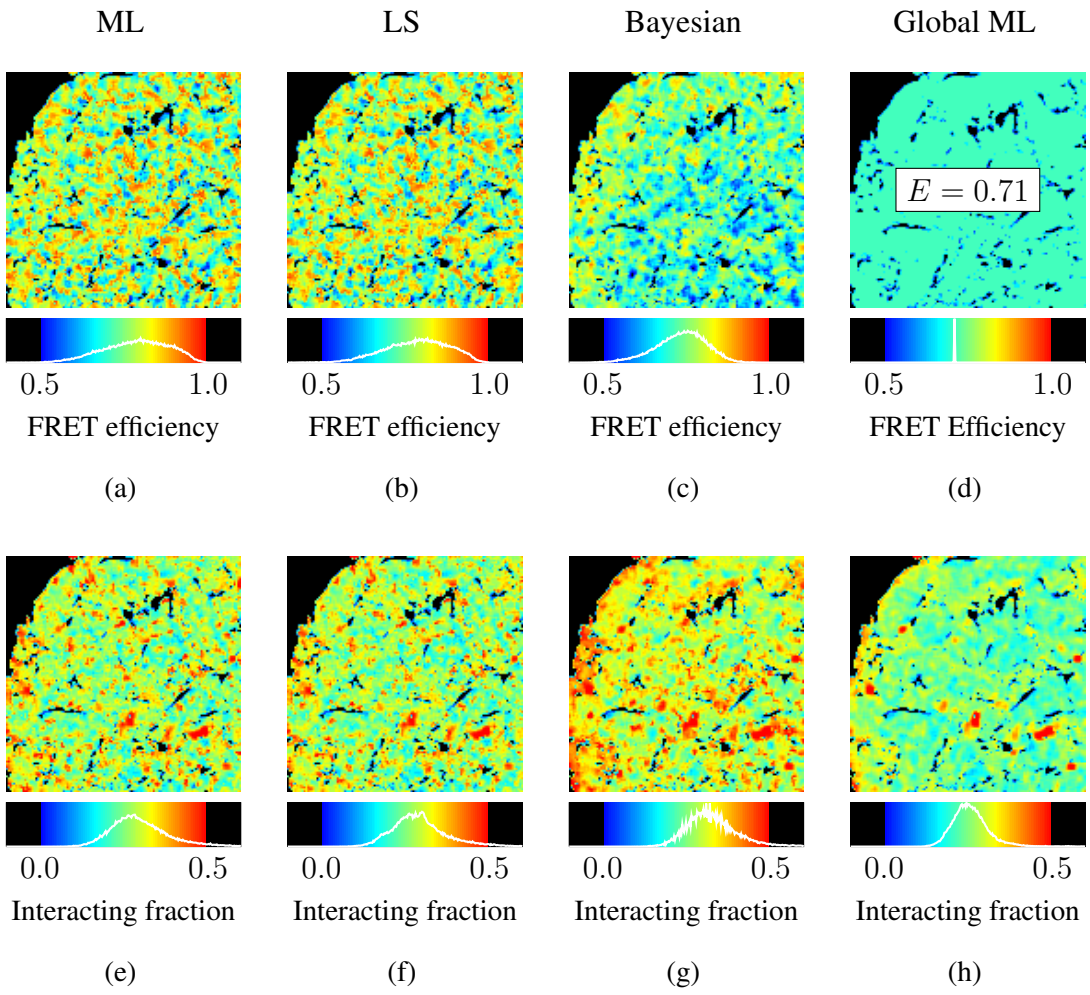


Figure 7.19: *FRET efficiency and interacting fraction estimation:* The FRET efficiency and interacting fraction maps obtained from the bi-exponential decay lifetime and initial amplitude estimates due to ML, LS, and Bayesian analysis, for the analysis of breast cancer tissue data. In (a), (b), (c), and (d), the FRET efficiency maps due to ML, LS, Bayesian analysis, and global ML analysis, respectively, and in (e), (f), (g), and (h) the corresponding interacting fraction maps for the different methods.

tissue core as suggested by the Bayesian analysis, is significantly broader than extent of the 7×7 spatial bin, and so is not likely to be an artefact of spatial binning.

The differences apparent in the interacting fraction and FRET efficiency maps due to ML and LS, due to global ML analysis, and due to Bayesian analysis are sufficiently pronounced that their respective parameter estimates may be used to support different biological conclusions. The remainder of this section is devoted to attempting to shed

light on why these differences between the Bayesian analysis estimates and ML, LS, and global ML analysis may exist, especially given that the bi-exponential estimates offered by the different analysis methods have been shown to be consistent with each other and with expectation using synthetic data in Section 7.1. It is essential that all possible reasons for the differences between the estimates of Bayesian analysis, global ML analysis, and those of ML and LS are considered; the purpose of the acquisition and analysis of time resolved data is to yield reliable decay parameter estimates from which reliable inference regarding the underlying biology may be made.

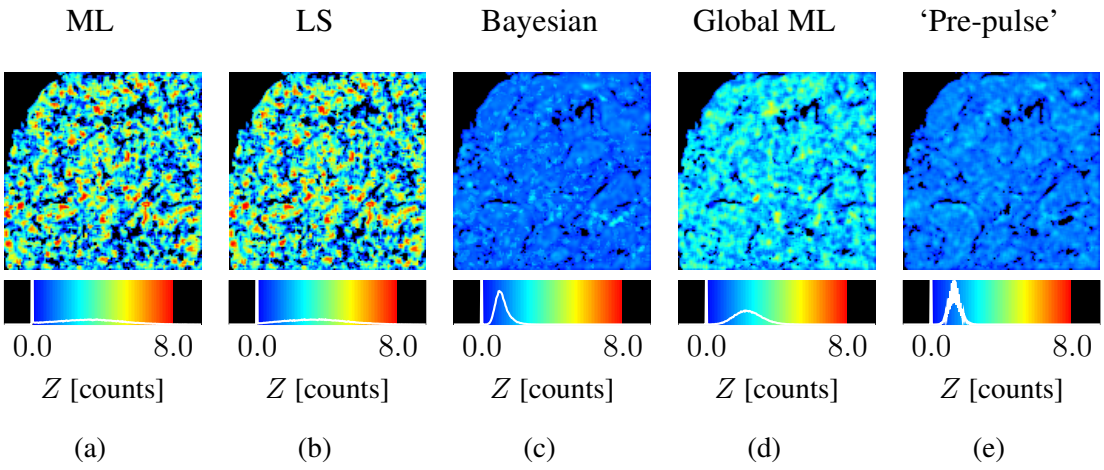


Figure 7.20: *Bi-exponential uniform background estimation:* The uniform background estimates due to ML, LS, Bayesian analysis, and global ML analysis, for the analysis of data acquired on imaging breast tissue. In (a), (b), (c), and (d), the estimated background maps due to ML, LS, Bayesian analysis, and global ML analysis, respectively.

In attempting to trace the source of the differences in the estimated FRET efficiency and interacting fraction due to ML and LS and due to Bayesian analysis, the plausibility of the uniform background estimates of the different methods is first considered. The uniform background maps due to the different analysis methods are shown in Fig. 7.20, along with the background estimates obtained on averaging the counts in the bins that precede the rise of the transient[1] at each pixel (i.e. the so called 'pre-pulse' portion of the measure-

[1] Although considerable effort has been expended in this work to develop a FLIM system model that properly accounts for photon counts that may be present in the 'pre-pulse' portion of the measurement interval as they may be due to slow fluorescence (i.e. the fluorescence may not have fully decayed during one repetition period) and not due to background, the breast tissue data analysed for this section was acquired with a repetition period of about 25 ns and the slow component of the fluorescence decay having a lifetime

ment interval); it is evident that the Bayesian estimation of the uniform background differs significantly from the estimation due to the other analysis methods, the Bayesian analysis estimates are distributed around a peak uniform background value of about 1.0 counts/bin but both ML and LS analysis are suggestive of a much higher background level, peaked between about 3.0 counts/bin and 3.5 counts/bin, and global ML analysis estimating a background of about 2.1 counts/bin. The background estimates obtained on averaging the counts in the 'pre-pulse' portion of the measurement interval are distributed around an average value of 1.35 counts/bin, and are in reasonably close agreement with those obtained using Bayesian analysis. Inspection of the time resolved data (not shown) confirms that the Bayesian estimated uniform background level is most consistent with experimental reality and that both ML and LS overestimate the background level considerably, and global ML analysis slightly less so. It would be difficult to conceive, for such low lifetimes and long excitation periods, of a reasonable explanation for background estimates that are consistently greater than those obtained on averaging the counts in the 'pre-pulse' portion of the measurement interval; it is plausible that the estimated background be less than the 'pre-pulse' determined background and that a slow decay component account for the remainder of the measured intensity in the 'pre-pulse' region.

In an effort to force the use of credible background estimates in ML analysis the following approach was adopted; for each pixel, the background estimate was determined from the so called 'pre-pulse' portion of the measurement interval and fixed for the subsequent estimation of the decay lifetimes and amplitudes. It is worthy of note that ML analysis with the background Z estimated from the 'pre-pulse' portion of the measurement interval offered strikingly different FRET efficiency and interacting fraction maps. The estimates obtained using ML analysis without 'pre-pulse' background estimation yield an average FRET efficiency and interacting fraction of 79% and 29% respectively, but the estimates obtained with 'pre-pulse' background estimation suggest a considerably lower average FRET efficiency of 67% and a higher average interacting fraction of 36%. To confuse matters yet further, global ML analysis with the background parameter fixed to the average (over the entire image) of the 'pre-pulse' estimated background values yielded lifetime estimates of 3.28 ns and 1.08 ns (i.e. a FRET efficiency of 67%), considerably different to the estimates of 3.13 ns and 0.90 ns (FRET efficiency of 71%) obtained with-

not exceeding about 3.5 ns, it is not unreasonable to assume that such counts falling in the 'pre-pulse' region are due to background.

out fixing the background parameter value.

Unfortunately, as all of the different analysis methods yield very different results it is neither possible to meaningfully compare the accuracy of the Bayesian estimates with those from the other analysis methods nor to make any strong claim regarding the validity of any of the results; it should be noted, though, that the developed Bayesian bi-exponential analysis makes use of all of the data in the measurement window, including that in the 'pre-pulse' region, and yields estimates for the background Z that are credible, whereas the established techniques do not. The lifetime and amplitude estimates that accompany the background estimates due to Bayesian analysis are, of course, sensitive to the estimated background. It is not unreasonable though to suspect that if, for whatever reason, the background is systematically over-estimated that this will have a bearing on the credibility of the accompanying lifetime and amplitude estimates.

Chapter 8

Bayesian simultaneous instrument and decay (SID) analysis

In this chapter the role of the instrument response function (IRF) on decay data analysis and parameter estimates is explored in detail. The IRF of a FLIM system is usually measured by recording reflected or scattered excitation light [106]. The IRF of most detectors is wavelength dependent [107] and, therefore, the IRF measured at the excitation wavelength may differ from that at the fluorescence emission wavelength. In [106], a method of recording the IRF at the emission wavelength, by imaging a quenched Rose Bengal solution having a lifetime of 16 ps, was demonstrated for a TCSPC FLIM system. However, it should be noted that, in [106], it was found that for common detectors measurement of the IRF using quenched Rose Bengal solution gave "practically the same results as with scattering". As the Bayesian SID (Simultaneous Instrument and Decay) analysis algorithm estimates an IRF from the fluorescence decay data, wavelength dependence is not an issue. The Bayesian SID algorithm may be useful when the IRF is not the same at each image pixel, such as may be the case for some widefield imaging applications, and in cases where the measurement of the IRF is not easy, for example, such as in endoscopic FLIM.

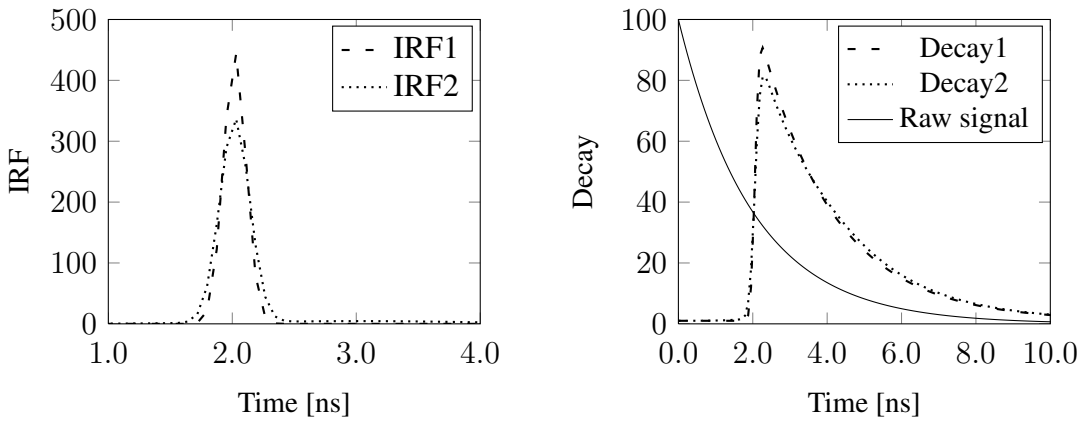
The IRF, whether somehow measured or estimated from decay data or otherwise, describes the likelihood that the experimental apparatus introduces a delay of a given duration between the emission of a decay photon and its subsequent detection. Although there are a few situations where the IRF can reasonably be ignored in decay analysis

(e.g. a narrow IRF and relatively broad bins), the examples and analyses presented in this chapter not only showcase Bayesian SID analysis but also serve to emphasize how critically important the IRF is to obtaining good decay parameter estimates regardless of the analysis technique employed, particularly in this work where parameter estimates are sought when data is often in short supply. The effects of using not quite the correct IRF on decay parameter estimates are demonstrated in Section 8.1 by means of a simple example. In Section 8.2 the Bayesian estimation of the IRF from decay data alone is investigated through three examples that simulate different IRFs, each posing a different challenge to the algorithm. An analysis of the application of the Bayesian SID algorithm to biological data, both from the viewpoint of a lifetime analysis and also as a means of IRF estimation at low counts, is presented in Section 8.3, before the advantages of and drawbacks inherent to Bayesian SID (and more generally to the approach of using an analytic approximation of the IRF in analysis) are considered in Section 8.4.

Throughout this chapter, in describing the Gaussian instrument response components, both the FWHM (full width at half maximum) and the standard deviation are presented, and are related by $\mathrm{FWHM} = 2\sqrt{2\ln 2}\,\sigma \approx 2.35\sigma$ for a Gaussian distribution of standard deviation σ.

8.1 Pitfalls of not using the correct instrument response

Is the signature of the experimental apparatus in the data really that important? It is instructive, somewhat alarming even, to observe what can happen to decay parameter estimates when not using the correct IRF in decay analysis. In this example, the simplest possible case of mono-exponential analysis is explored; two synthetically generated images, both containing at each pixel identical mono-exponential decays, but simulated to be perturbed by slightly different instruments are analysed (using ML). One image (Image1) simulates an IRF (IRF1) consisting of a single Gaussian and the other image (Image2) simulates a slightly different IRF (IRF2) comprising a slightly wider Gaussian and a tail to its right; such an instrument having the effect of spreading more arrival times towards later bins (in the measurement interval). Although on visual inspection (Fig. 8.1) neither IRF1 and IRF2 nor their respective perturbed decay signals appear to differ significantly, the biasing of lifetime estimates obtained when using the wrong IRF for decay analysis is clearly apparent (Fig. 8.2).

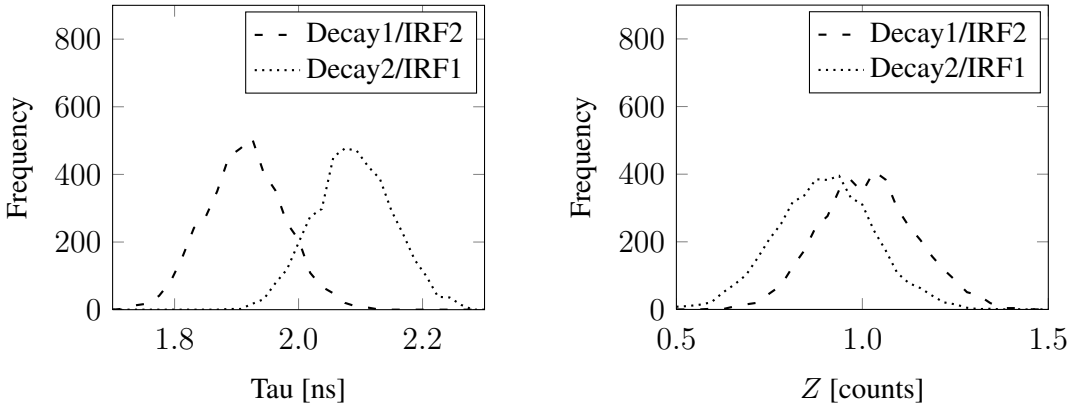


(a) IRFs

(b) Decays

Figure 8.1: *Slightly different instruments, slightly different decays:* In (a) the two synthetically generated IRFs used for the generation of Image1 and Image2; IRF1 being composed of a single Gaussian of FWHM 150 ps (standard deviation 64 ps), and IRF2 being composed of a dominant component (ratio 10:1) that is slightly wider than IRF1 and an additional component to simulate the tail to the right (being retarded by a further 1.0 ns and having FWHM 1500 ps (standard deviation 640 ps). In (b) the mono-exponential signal of lifetime $\tau = 2.0\mathrm{ns}$, having constant background $Z = 1.0$ counts/bin and initial amplitude $A = 100.0$ counts, and the ideal (i.e. noise-free) decay curves resulting from the convolution of IRF1 and IRF2. (Note that the time scales in (a) and (b) differ in order that the differences between IRF1 and IRF2 can be more easily seen.)

It is not surprising that the parameter estimates (Fig. 8.2), in this simple example the constant background level and the mono-exponential lifetime, are sensitive to the IRF used in the decay analysis. The lifetime estimates obtained if using IRF1 in the analysis of IRF2 decay data are systematically overestimated while the background level is underestimated. Intuitively, this seems reasonable and is fairly easily explained as follows; when using IRF1 for the analysis of IRF2 decay data the presence of photon counts in later bins is misinterpreted as being due to a longer decay lifetime, and as a consequence of more photon counts being attributed to a slow decay then less counts are attributed to background. A similar argument can be applied to explain the underestimation of the decay lifetime and the overestimation of the background level if using IRF2 in the analysis of IRF1 decay data. Reassuringly, the lifetime and background estimates are as should be expected (not shown) when the images are analysed with the appropriate IRFs.



(a) Lifetime estimates

(b) Background estimates

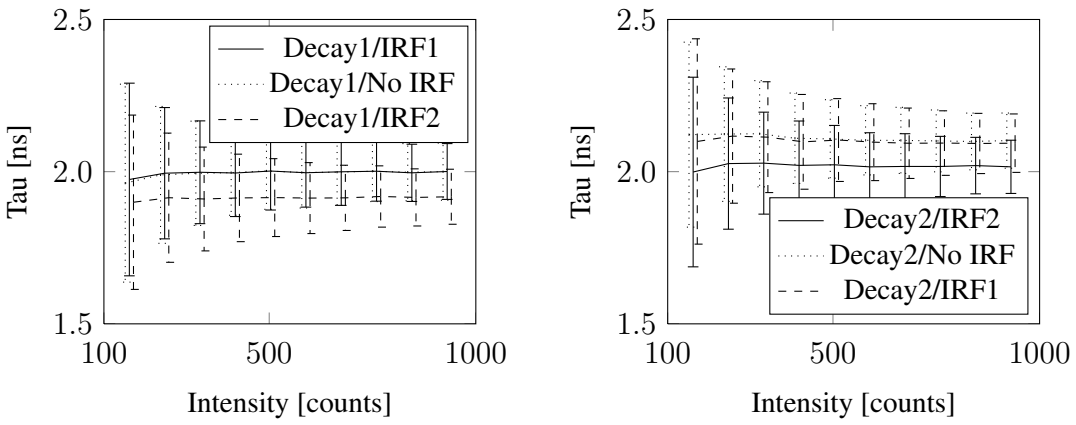
Figure 8.2: *Slightly incorrect analysis, very misleading decay estimates:* In (a) the lifetime estimates resulting from analysis of Image1 (i.e. Decay1 data) using IRF2, and the analysis of Image2 using IRF1 in analysis. It is clear that the analysis of Decay1 data with IRF2 shows systematic underestimation of the lifetime parameter from the true value of 2.0 ns. There is systematic overestimation of the lifetime when analysing Decay2 data with IRF1. In (b) the estimated constant background, which is overestimated when using IRF2 in the analysis of Decay1 data, and underestimated when using IRF1 in the analysis of Decay2 data. With the incorporation of Poisson noise, for both images each image pixel contains about 2800 photon counts.

The data isn't that good anyway; is the signature of the experimental apparatus still important in the low count regime? The example presented above illustrates clearly that using an incorrect IRF in data analysis can have a significant impact on the reliability of decay parameter estimates, and could potentially even distort the overall findings of an experiment or study should the IRF used in analysis be crudely different from that of the experimental system on which the data was acquired. However, the biased estimates presented in Fig. 8.2 were obtained on the analysis of data containing around 2800 photon counts (per image pixel); not high count data but certainly not very low count data for a mono-exponential analysis. The remainder of this section is devoted to considering the importance of the IRF in the analysis of low photon count data; in particular to investigate whether there comes a point[1] where the uncertainty in any decay estimates is so signifi-

[1]Even the approximate total photon count at which such a point may be will obviously differ greatly from case to case; for example, more photon counts are required for the accurate mono-exponential analysis of data acquired with a significant background than with only a negligible background, and many more still for a bi-exponential analysis.

cant that it is not necessary to be overly concerned by the more intricate details such as the accuracy of the IRF used for analysis. The example above is now extended in order to provide insight as to whether the IRF remains important in the low photon count regime. Data having between about 100 and 1000 total counts were simulated using the same two IRFs (Fig. 8.1) and a mono-exponential decay of lifetime 2.0 ns; as previously data incorporating IRF1 were termed Decay1, and that in which IRF2 is simulated were termed Decay2. In addition to analysis with the correct and the incorrect IRF, the low count data was also analysed using no IRF at all. On inspecting Fig. 8.3, it is immediately evident that the biasing of parameter estimates when analysing decay data with the wrong IRF does not diminish with decreasing total photon counts, the decay lifetime being systematically underestimated when analysing Decay1 data with IRF2 and overestimated when analysing Decay2 data with IRF1. However, it also apparent that the lifetime estimates obtained from Decay1 data are almost no different when no IRF is used as compared to when the correct IRF (IRF1) is employed in data analysis; this is accounted for by IRF1 being symmetrical and relatively narrow compared to the bin width and thereby not spreading the recorded photon arrival times greatly beyond the bins into which, without any spreading, they would be anticipated anyway. The same is not true of the analysis of Decay2 data with no IRF, however, with the decay lifetime being systematically overestimated; again, intuitively, this is consistent with expectations as the Decay2 data shall contain more photon counts in higher bins due to IRF2 having a tail to the right and analysis without an IRF does not account for this.

Clearly, as this example, which was constructed for the simplest possible case of a mono-exponential decay and instrument responses that can be described perfectly analytically, demonstrates, the importance of the IRF in analysis should not be underestimated. In reality, with potentially multi-exponential decay signals and a less analytically perfect instrument response, analysis problems are unlikely to diminish. With so many variables inherent to a data acquisition system and the fluorescence decay, there is no definitive answer as to whether or how important the IRF (or an accurate IRF measurement) is in either the high or the low count regime. One may be fortunate and obtain perfectly reliable estimates (e.g. decay lifetime(s)) without using the correct IRF in analysis; however, just as one would be unlikely to have complete confidence in the results of an analysis without having confidence in the data acquisition and the experimental protocol as a whole, one should be wary that the integrity of any estimates may also be threatened by the IRF used



(a) Decay1 analysis

(b) Decay2 analysis

Figure 8.3: *Does the IRF influence parameter estimates greatly when total counts are meagre?:* In (a) lifetime estimates resulting from the analysis of data simulated to incorporate IRF1 (i.e. Decay1 data) when using the correct IRF (IRF1), an incorrect IRF (IRF2), and no IRF at all, for low count data having between about 100 and 1000 total photon counts. In (b) lifetime estimates from the analysis of data simulated with IRF2 (i.e. Decay2 data) when using the correct IRF (IRF2), an incorrect IRF (IRF1), and no IRF at all.

in analysis.

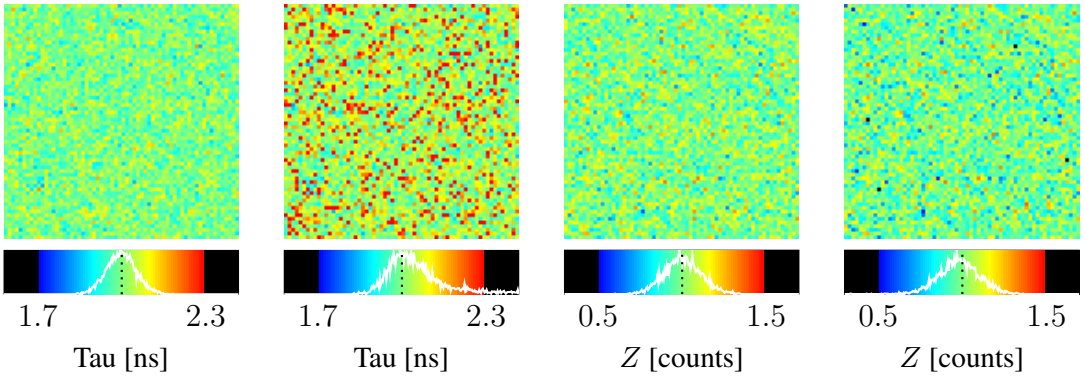
8.2 Synthetic data examples

Suppose that the images analysed for the example in Section 8.1 were to contain data acquired on real systems and that, for whatever reason, a lifetime analysis is required but no information regarding the IRFs of the acquisition systems are available. Would it be possible to proceed with a *meaningful* analysis? How could one best approach this problem if all of the available algorithms require an IRF for analysis? Assume, for the sake of argument, that the data had been previously collected but that fortunately access to the acquisition system remains possible; one could with knowledge of the system in its *current* state attempt to somehow approximate the IRF of the earlier state when the data were acquired. Otherwise, if possible, one could measure the IRF and hope that it does not differ significantly from that that would have been obtained at the time of data acquisition and that will be inherent to the decay data. Should access to the acquisition

system no longer be possible things are a little more difficult; one could somehow approximate an IRF for use in analysis based on the IRF of a similar system, or, one could even perform an analysis without any IRF in the hope that its influence on the data (and any subsequent parameter estimates) is minimal. Some of these workarounds may even yield useful decay estimates. However, all of them require additional (sometimes considerable) effort and preferably access to the system with which the data were acquired. Crucially, though, none make use of the most readily available and potentially reliable source of information regarding the IRF at the time the data was acquired; the decay data itself. Of course, the decay data contains information about both the decay process and the IRF of the system on which the data were acquired. The Bayesian SID analysis developed in Chapter 5 offers a means of extracting parameter estimates that describe both the decay and the approximated IRF. Lifetime estimation would be possible using the Bayesian SID algorithm in a case such as that described here (although, in this case, the IRF parameter estimates would be incidental as it is a lifetime analysis that is desired), without resorting to any of the methods required for the workarounds suggested.

The decay parameter estimates resulting from two slightly different approaches that use the Bayesian SID algorithm for the analysis of Image1 (Section 8.1) are shown in Fig. 8.4. The first, and most commonly used, approach employs the Bayesian SID algorithm as a pre-step, assuming that the IRF (and implicitly the decay order) is uniform across an image, to obtain the optimal IRF approximation from the large data set produced on summing the data of each image pixel, for use in a subsequent standard Bayesian decay analysis. The second approach is to apply the Bayesian SID algorithm independently at each image pixel, thereby making no assumptions regarding uniformity of the IRF across an image, and to estimate the optimal decay and IRF parameter values at each pixel.

On inspecting the lifetime and background estimate images in Fig. 8.4 and the estimated lifetime distributions in Fig. 8.5, it is clear that the decay parameter estimates are superior when Bayesian SID analysis is employed as a pre-step to a conventional decay analysis that uses the optimal IRF approximation; the lifetime estimate distribution being centered around an average of 2.00 ns with standard deviation 0.05 ns. The lifetime estimates for application of the Bayesian SID algorithm at each image pixel are centered around an average of 2.04 ns with standard deviation 0.08 ns. Of course, it is worth bearing in mind that the estimates are just that; in demanding of the Bayesian SID algorithm the estimation of more model parameters (both decay parameters plus IRF parameters),



(a) Decay analysis (b) SID analysis (c) Decay analysis (d) SID analysis

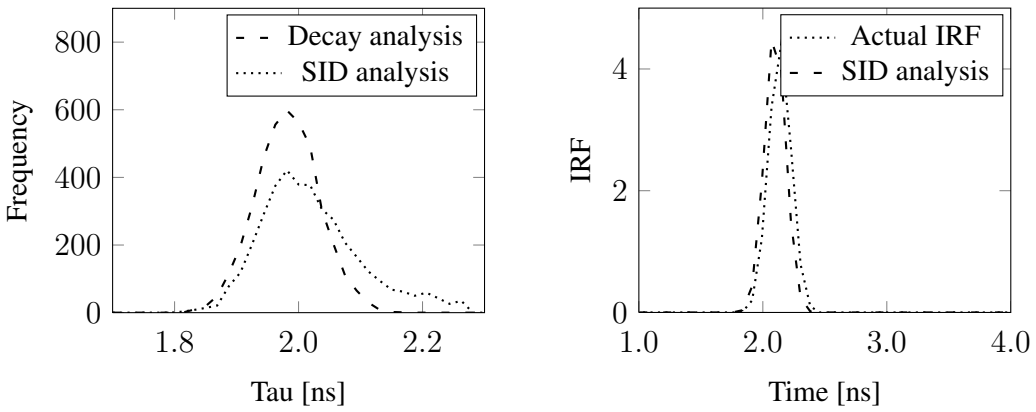
Figure 8.4: *Bayesian SID analysis in the absence of any IRF details:* The decay parameter estimates obtained using the Bayesian SID algorithm without any information regarding the IRF for the analysis of Image1 of Section 8.1 (having about 2800 total photon counts at each image pixel). In (a) and (c) the lifetime and constant background images respectively obtained using Bayesian mono-exponential analysis having as a pre-step determined the optimal IRF parameter values using the Bayesian SID algorithm on the single data set produced by summing the data of all image pixels. In (b) and (d) the lifetime and background estimates obtained on the application of the Bayesian SID algorithm at each image pixel.

with no additional information for analysis, it should not be surprising that the decay parameter estimates are not quite as good as those from decay analysis alone with an optimal IRF approximation. When performing Bayesian SID analysis as a pre-step to a conventional Bayesian decay analysis, the analysis yields decay parameter estimates *given* the data and *given* the optimal IRF approximation; on the application of the Bayesian SID algorithm alone the analysis yields decay parameter estimates *and* IRF parameter estimates *given* the data.

The remainder of this chapter explores in greater detail the performance of the Bayesian SID algorithm, exploring the results of analyses from a variety of synthetic IRFs and also the results of analysis of real biological data.

8.2.1 Example 1: Single Gaussian

In this simple example the performance of the Bayesian SID algorithm is studied using synthetic mono-exponential decay (of lifetime 2.0 ns) data incorporating an IRF com-



(a) Bayesian lifetime estimation

(b) Bayesian IRF estimation

Figure 8.5: *Bayesian SID analysis, two different approaches:* Analysis of Image1 of Section 8.1 (having about 2800 total photon counts at each image pixel) comprising a single Gaussian IRF and a mono-exponential decay of lifetime 2.0 ns at each image pixel. In (a) the lifetime estimates resulting from Bayesian SID analysis applied to each image pixel and the estimates from a conventional Bayesian decay analysis using the optimal IRF approximation (as determined by the Bayesian SID algorithm as a pre-step). In (b) the actual synthetic IRF and the Bayesian SID approximation estimated from high count data (about 10^7 total counts).

prised of a single Gaussian component of FWHM 212 ps (standard deviation 90 ps). In order to observe almost the simplest possible case, no constant background is incorporated in the simulated data so as to avoid any additional complications that this may cause.

On visual inspection of the IRF (Fig. 8.6), it is clear that it is symmetrical and will most likely introduce a delay of between about 2.0 ns and 2.5 ns between the recorded data and any fluorescence decay photons, with the majority of photons being delayed by the duration at which the IRF peak is found (closer examination of the numerical data shows this to be around 2.10 ns). It is also worthwhile studying the typical time resolved data sets as shown in Fig. 8.6 and considering what information about any features of the IRF are readily apparent. Inspecting the high count data set, it can be seen that the rise in the transient does not occur instantaneously, happening between about 2.0 ns and 2.1 ns; it is easy to deduce therefore that the IRF inherent to the data does not merely introduce a delay between decay photons and their recorded counterparts but also spreads them (i.e. introduces some uncertainty into the data recorded). Very little can be gleaned from observing the low count data set (having about 100 total photon counts), aside from

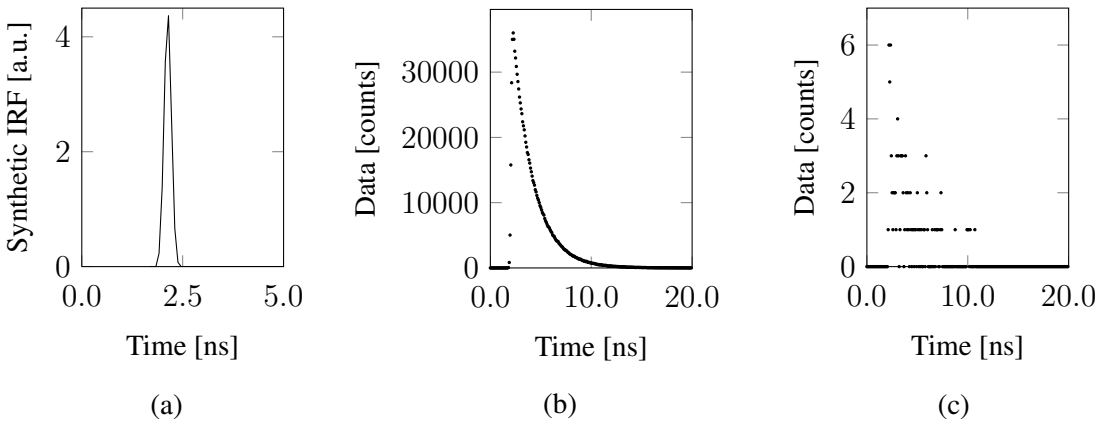
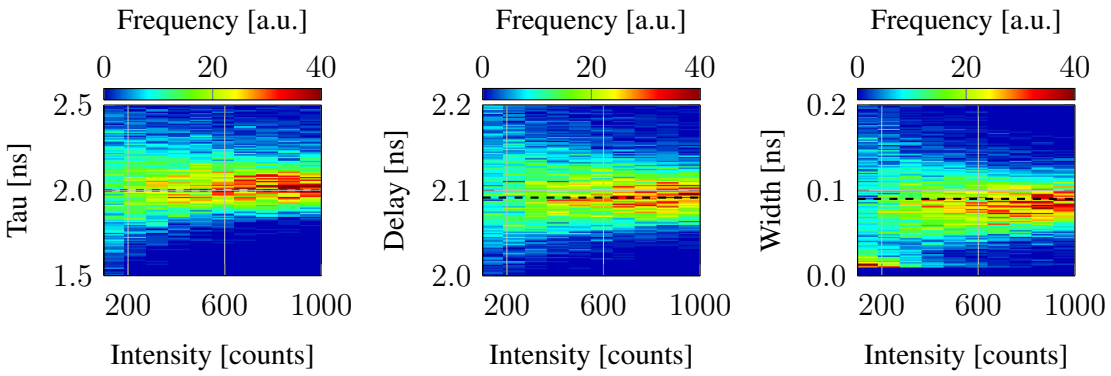


Figure 8.6: *A single Gaussian IRF:* In (a) the synthetic IRF composed of a single Gaussian distribution of FWHM 212 ps (standard deviation 90 ps). In (b) and (c) typical time resolved data from the convolution of such an IRF with a mono-exponential decay of lifetime 2.0 ns, having about 10^7 and 100 total photon counts respectively.

perhaps that the IRF introduces a delay into the data and only the most optimistic would even attempt to quantify the delay on the basis of such data.

The Bayesian SID algorithm decay lifetime and IRF parameter estimates are shown in Fig. 8.7 for data having between about 100 and 1000 total counts. As expected, the parameter estimates improve as the total photon count increases, as is particularly evident for the IRF width parameter, the Bayesian SID algorithm defaulting to the minimum width value permitted by the implementation, most of the time, until the intensity exceeds about 300 total counts. Inspecting the distribution of Bayesian SID decay lifetime estimates (Fig. 8.7), although most of the estimates are (at least within a reasonable uncertainty given the low count nature of the data) close to the true decay lifetime of 2.0 ns, it is also evident that in some cases the decay lifetime is significantly overestimated. The tendency for the Bayesian SID algorithm to sometimes overestimate the decay lifetime is more readily observed in Fig. 8.8, which shows the estimated lifetime distributions for the Bayesian SID algorithm and for Bayesian decay analysis (using the optimal IRF parameter values as determined by Bayesian SID analysis with very high count data (10^7 total photon counts)) for data containing about 1000 total photon counts; the estimates from Bayesian decay analysis do not exceed about 2.2 ns whereas the distribution of Bayesian SID estimates has a substantial tail extending to lifetimes of about 2.5 ns. The introduction of the IRF width and delay parameters into the analysis model, it seems, has



(a) Decay lifetime (b) IRF delay (c) IRF width

Figure 8.7: *Bayesian SID analysis IRF and decay lifetime estimation with increasing total photon count (I):* In (a) the Bayesian SID analysis estimates of the mono-exponential decay lifetime parameter for increasing total photon count. In (b) and (c) the corresponding (single Gaussian component) IRF delay and width parameter estimates respectively. The dashed black lines show the true parameter values in each plot.

introduced unwanted correlations between the model parameters and has lead to the decay lifetime being significantly overestimated at times. It is not easy to discern a correlation between, say, an overestimated lifetime and an underestimated IRF parameter width. The relationships between the decay lifetime and the decay background and the IRF delay and width parameters are shown in Fig. 8.8; the background level estimates are roughly symmetrical around the true decay lifetime, as are the IRF delay parameter estimates. However, it does appear (Fig. 8.8) that if the decay lifetime is overestimated it is likely that the IRF width is underestimated, this being particularly so for estimated lifetimes longer than about 2.1 ns.

Of course, fortunately here it is considerably easier to inspect for any unwanted correlations between decay and IRF parameter estimates as in this case the true parameter values are known. Usually one wouldn't be equipped with such knowledge, or equally such a detailed analysis of parameter estimates may not be undertaken and correlations not detected. It must be stressed that such unwanted correlations between model parameters cannot be disregarded as merely troublesome features in the estimates distribution; in a real experiment and analysis such features may lead to an incorrect interpretation. The existence of such unwanted correlations certainly reinforce the case for a reasonably strong prior on the IRF parameters should any reliable information from, say, a measured

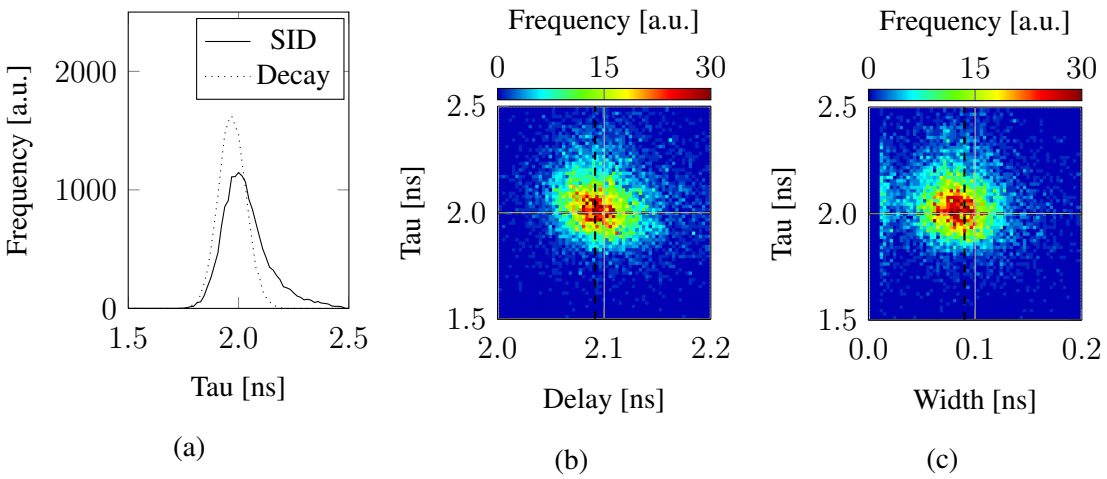


Figure 8.8: *Bayesian analysis decay lifetime estimation with increasing total photon count (II):* In (a) the estimated lifetime distributions for the Bayesian SID algorithm and for Bayesian decay analysis (using the optimal IRF parameter values as determined by Bayesian SID analysis with very high count data (10^7 total photon counts)) for data containing about 1000 total photon counts. In (b) the correlations between the Bayesian SID estimated decay lifetime and IRF delay parameter, and in (c) the same for the IRF width parameter for the mono-exponential decay data at low photon counts.

IRF be available. However, particularly on application of the Bayesian SID algorithm for the analysis of low count data, should any evidence be available regarding any characteristics of the instrument, such as the likely IRF width (or even the width of say a dominant component), then it would certainly be reasonable to incorporate that in a strong prior distribution over the IRF parameters.

8.2.2 Example 2: Two resolvable Gaussian IRF components

In this example the performance of the Bayesian SID algorithm is studied using synthetic mono-exponential decay (of lifetime 2.0 ns) data incorporating an IRF comprised of two Gaussian components in a 4:1 ratio; both components being of width 90 ps (FWHM 212 ps) with the minor component being retarded by 1 ns relative to the dominant component. It is clear merely on visual inspection of the IRF (Fig. 8.9) that it is composed of two distinct (and easily resolved) components. Additionally, it is worthwhile studying the typical time resolved data sets as shown in Fig. 8.9 and considering what information about any features of the IRF are readily apparent. On visual inspection of the high count data set, it would be reasonable to surmise that the IRF is comprised of *at least* two distinct components (assuming an exponential decay signal of course). It is not possible, however, to reach such a conclusion on inspecting the low count data set (having about 1000 photon counts in total).

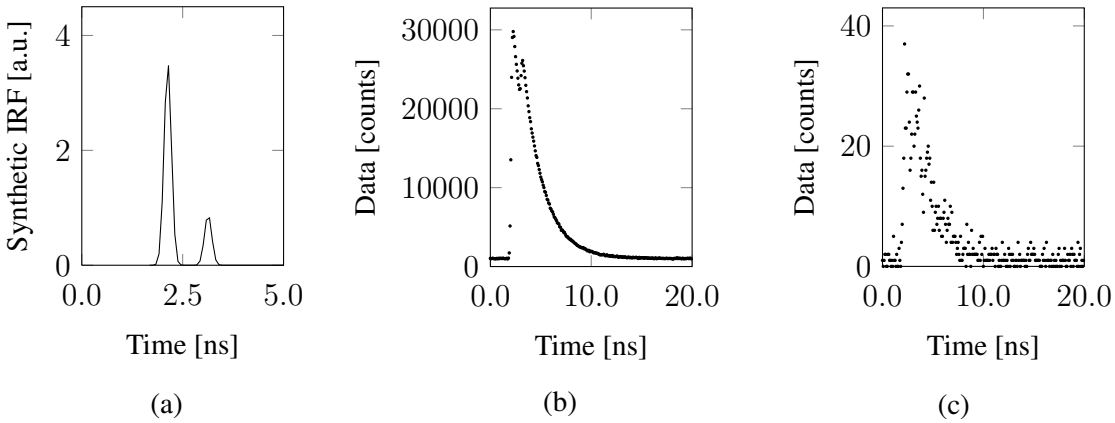


Figure 8.9: *Two resolvable Gaussians:* In (a) the synthetic IRF composed of two Gaussian distributions in a 4:1 ratio, both components being of width 90 ps (FWHM 212 ps) and the minor component being delayed by 1 ns relative to the dominant component. In (b) and (c) typical time resolved data from the convolution of such an IRF with a mono-exponential decay of lifetime 2.0 ns, having total counts of about 10^7 and 1000 respectively.

The IRF parameter estimates offered by Bayesian SID analysis of an image where each pixel has data containing about 1000 total counts and another image with each pixel having data sets of about 10000 total counts, are shown in Fig. 8.10. On inspecting the IRF parameter estimate images for data having about 1000 total photon counts it is evident

that although the estimates that define the dominant component are reasonably well defined, and that at the majority of pixels the Bayesian SID analysis estimates do indicate an IRF having two distinct components, for such low count data given the relative complexity of the simulated instrument the IRF parameter estimates may not be sufficiently robust to be of practical use; they do however offer a greater insight than does visual inspection of the decay data. The estimates for data containing about 10000 total counts, though, are tightly distributed other than the width of the minor component; the weighting of the dominant component is centered around a value of 0.797 with standard deviation 0.043 (true value 0.8), the delay and width of the dominant component are centered around average values of 2.091 ns (optimal value 2.09 ns) and 0.084 ns with standard deviation 0.009 ns and 0.008 ns respectively, and the delay and width of the minor component are centered around the values 3.081 ns (optimal value 3.09 ns) and 0.079 ns with standard deviation 0.082 ns and 0.065 ns respectively. No such statistics are presented for the low count data estimated parameter value distributions as they display multi-modal features (other than the delay of the dominant component). It is clear though that even with data containing around 10000 total counts that the width of the minor component is poorly estimated; however, it should also be noted that the accompanying lifetime estimates (Fig. 8.11) are well defined around the true value and compare reasonably well the estimates obtained from conventional Bayesian mono-exponential decay analysis alone using the optimal IRF approximation.

The Bayesian SID estimated decay lifetime images are shown in Fig. 8.11, as are the lifetime images obtained from a Bayesian mono-exponential decay analysis (using the optimal Bayesian SID estimated IRF, as obtained from a very high count data set having about 10^7 total photon counts) of the same data. The cleaner lifetime images illustrate that the estimates obtained from conventional Bayesian decay analysis using the optimal IRF approximation are superior to those from Bayesian SID analysis; this should of course be expected as five fewer model parameters are estimated from the same data. The more tightly distributed estimates are especially apparent for the 1000 total counts per pixel image, and less so for the higher count data image (each pixel having about 10000 total counts). For the 10000 total counts data, the lifetime estimates of both conventional Bayesian mono-exponential decay analysis and of Bayesian SID analysis are centered around the true value of 2.0 ns, being distributed with standard deviation 0.038 ns for Bayesian SID analysis and 0.024 ns for conventional decay analysis using the optimal

IRF approximation.

The Bayesian SID decay lifetime estimates are shown with those obtained from mono-exponential Bayesian decay analysis for intensities between about 1000 and 10000 total photon counts in Fig. 8.12. It is evident that below about 3000 total counts the Bayesian SID estimated lifetime may be prone to overestimation; further analysis may well show that there is a correlation between an overestimated lifetime and one (or more) of the estimated IRF parameters. As discussed in the previous example, should sufficient evidence exist to justify the use of a strong prior over (at least some of) the IRF parameters it may be possible to reduce unwanted correlations; regardless, caution should be exercised, especially should no evidence be available to support the use of a strong prior on the IRF parameters, and it should be acknowledged that unwanted correlations between some model parameters may exist.

It is also evident (Fig. 8.13) that the parameter estimates that define the minor IRF component are inferior to those of the dominant component; again, this is unsurprising as the signature of the minor component in the decay data is less pronounced and such subtle features may be far more difficult for even the most sophisticated analysis algorithms to capture, especially at low counts where noise in the data has a greater influence.

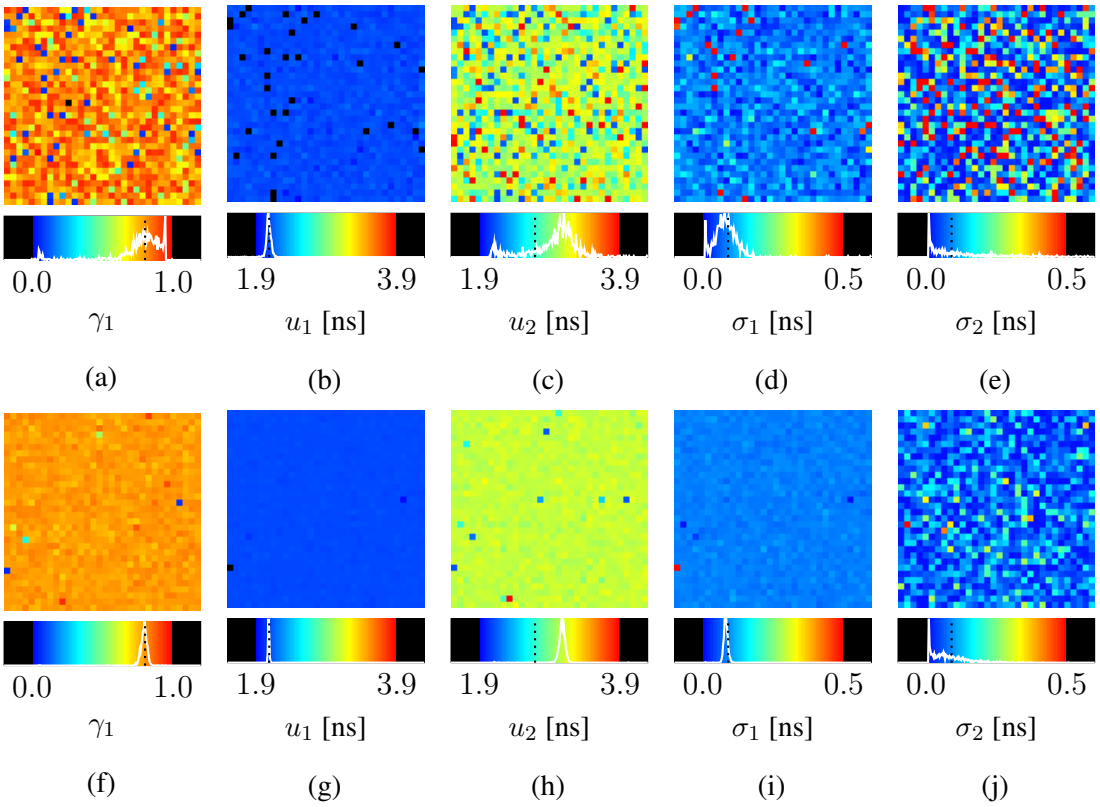
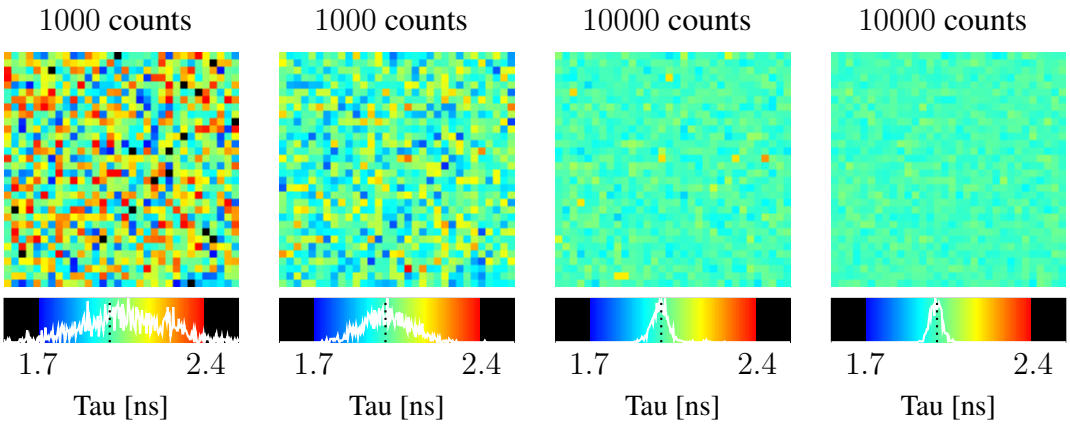
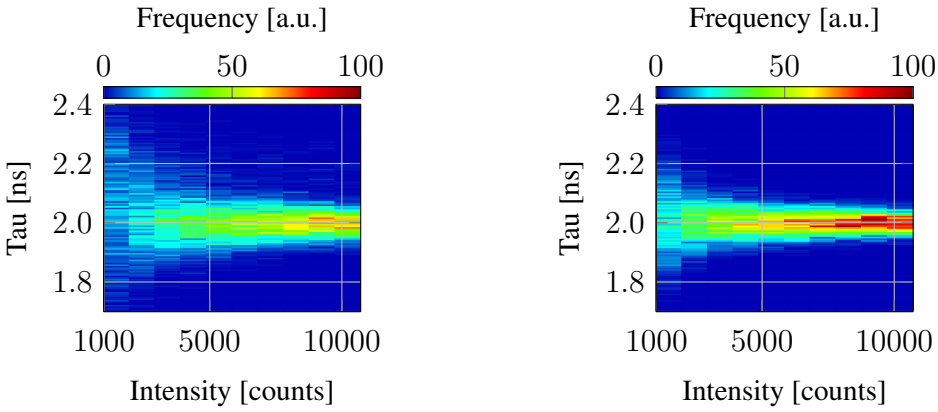


Figure 8.10: *Bayesian SID analysis for a two-component Gaussian IRF approximation:* Two 32×32 pixel intensity images, one with each pixel having a total photon count of about 1000 photon counts and the other with each pixel having about 10000 counts were analysed with Bayesian SID algorithm, the IRF parameter estimate images for the two-component IRF approximation are displayed. For the low count image, in (a) the estimated weight of the dominant component (should be 0.8), in (b) the delay parameter u_1 of the dominant component, in (c) the delay parameter u_2 of the minor component, in (d) the width parameter σ_1 of the dominant component, and in (e) the width parameter σ_2 respectively that describe the Gaussian IRF approximation. In (f), (g), (h), (i), and (j), the corresponding estimates for the high photon count image.



(a) SID analysis (b) Decay analysis (c) SID analysis (d) Decay analysis

Figure 8.11: *Bayesian SID analysis and Bayesian decay analysis lifetime estimates (I):* In (a) the estimated mono-exponential decay lifetime image from Bayesian SID analysis of synthetic image containing mono-exponential decay (lifetime 2.0 ns) data having about 1000 total counts at each of the 32×32 pixels, and in (b) the lifetime image produced from conventional Bayesian decay analysis of the same image. In (c) and (d) the Bayesian SID estimated lifetime and the Bayesian decay analysis lifetime images respectively, produced from the analysis of data having about 10000 total counts at each pixel.



(a) SID analysis (b) Decay analysis

Figure 8.12: *Bayesian SID analysis and Bayesian decay analysis lifetime estimates (II):* In (a) the Bayesian SID analysis estimates of the mono-exponential decay lifetime parameter for increasing intensity between about 1000 and 10000 total photon counts, and in (b) the corresponding Bayesian decay analysis lifetime estimates obtained using the optimal IRF approximation.

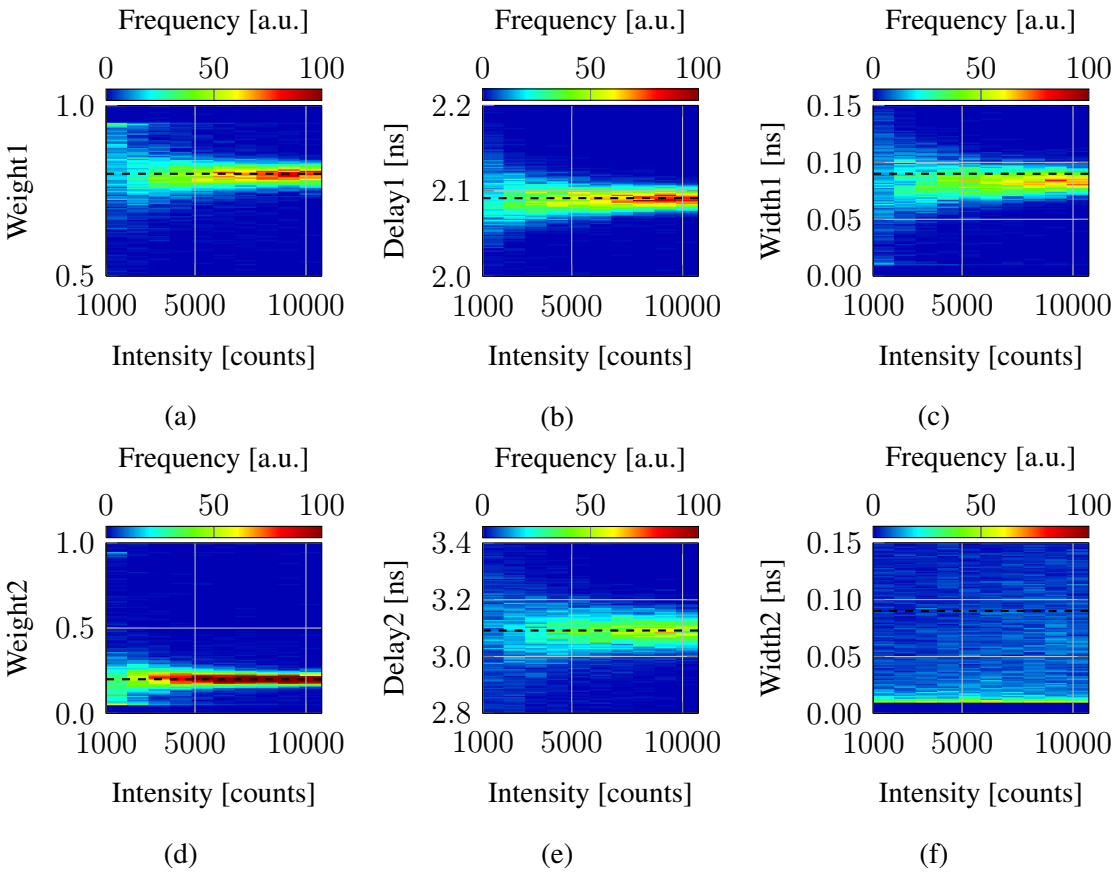
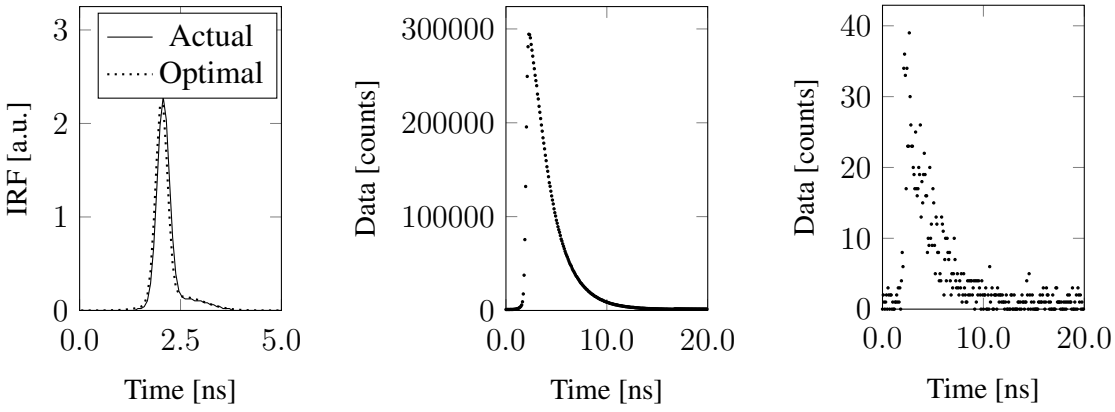


Figure 8.13: *Bayesian estimation of IRF parameters:* The Bayesian SID estimates of the two component IRF approximation parameters assuming a mono-exponential decay. In (a), (b), and (c), the parameter estimates that define the dominant component of the synthetic IRF, and in (d), (e), and (f), the parameter estimates that define the minor component.

8.2.3 Example 3: A tail to right

In this example the simulated IRF is composed of two Gaussian components in a 4:1 ratio; the dominant component being of FWHM 495 ps (standard deviation 210 ps) with the minor component being retarded by 0.5 ns relative to the dominant component and being of FWHM 1994 ps (standard deviation 846 ps) in order to construct the heavy tail to the right; it should also be noted that the minor component extends a little to the left of the dominant component thereby increasing the likelihood of early photon arrival times being recorded. As previously, the Bayesian SID algorithm is studied using synthetic mono-exponential decay (of lifetime 2.0 ns) data.

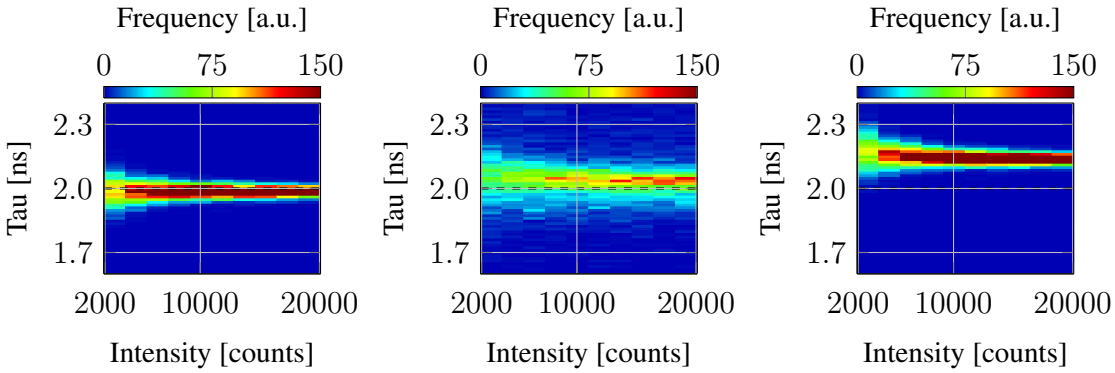


(a) IRF (b) 1000 counts (c) 10^7 counts

Figure 8.14: *An IRF with a heavy tail to the right:* In (a) the synthetic IRF composed of two Gaussian distributions in a 4:1 ratio, the dominant component being of FWHM 495 ps (standard deviation 210 ps) with the minor component being retarded by 0.5 ns relative to the dominant component and being of FWHM 1994 ps (standard deviation 846 ps). In (b) and (c) typical time resolved data from the convolution of such an IRF with a mono-exponential decay of lifetime 2.0 ns, having total counts of about 10^7 and 1000 respectively.

It is clear merely on visual inspection of the IRF (Fig. 8.14) that it is composed *at least* two overlapping components and that there is is significant likelihood of photons being delayed by a duration greater than that which would be attributed to the dominant component. However, in this example it is far more difficult to extract as much information about the IRF from the typical decay data, aside from it being clear that most photons are delayed by more than about 2.0 ns and there is likely also to be some spreading of photon

arrival times as evidenced by the non-instantaneous rise of the transient data. More subtly, on comparing the high count data set with those of the previous examples it can be seen that the transient rises more slowly than in the (Fig. 8.6, Fig. 8.9); this is a consequence of the minor component extending to the left of the narrower dominant component and thereby promoting the likelihood of recording photons a little earlier than those delayed by the duration corresponding to the dominant peak.



(a) Bayesian decay analysis (b) Bayesian SID analysis (c) ML with no IRF

Figure 8.15: *Bayesian and ML decay analysis lifetime estimates:* For increasing intensity between about 2000 and 20000 total photon counts, in (a) the lifetime estimates from a standard Bayesian mono-exponential decay analysis using the optimal IRF approximation (as determined by Bayesian SID with data containing about 10^7 total photon counts), in (b) the Bayesian SID analysis estimates of the mono-exponential decay lifetime parameter , and in (c) the lifetime estimates obtained from a ML analysis without using an IRF.

The mono-exponential decay lifetime estimates are shown for increasing intensities in Fig. 8.15 for standard Bayesian decay analysis (employing the optimal IRF approximation as determined by Bayesian SID using a data set containing more than 10^7 total photon counts), for Bayesian SID analysis, and for ML where no IRF has been used in the analysis (i.e. the same conditions under which the Bayesian SID algorithm operates). As would be expected at such intensities, the decay lifetime estimates obtained using standard Bayesian decay analysis with an optimal IRF approximation are tightly distributed around the true lifetime of 2.0 ns. It is clearly apparent that the Bayesian SID decay lifetime estimates are far more widely distributed than obtained from the standard Bayesian decay analysis; this is not surprising given that the parameters required to describe a two-component Gaussian

IRF approximation are also estimated along with the decay parameters. It is also evident that the Bayesian SID algorithm tends to systematically overestimate the decay lifetime, the mode of the distributions being at a lifetime slightly longer (at about 2.03 ns) than the true value of 2.0 ns even at an intensity of about 20000 total photon counts. Of course it would be desirable that the Bayesian SID algorithm (or any other analysis technique for that matter) would be able to extract extremely accurate (perfect) parameter estimates from low count data incorporating such an IRF. It is necessary to be realistic as to what can be achieved though; this is certainly a more challenging case than the previous examples, the later photon arrival times actually due to the heavy tail to the right of the IRF being attributed to a slower decay process (i.e. a longer lifetime). It should be noted that, even at high counts conventional ML analysis is unable to provide accurate lifetime estimates unless the measured IRF is provided; the lifetime estimates obtained from a ML analysis that does not use an IRF are prone to a significant systematic overestimation, the mode of the distribution being close to a value of 2.14 ns for all intensities investigated.

The dominant IRF component was estimated to be of width 200 ps (true value of 210 ps) and weighting 0.798 (true value of 0.8), with the minor component having a width of 928 ps (true value 846 ps) and being retarded by 0.57 ns relative to the dominant component (true value 0.50 ns). Intuitively, one might expect the overestimation of both the width of the minor IRF component and its delay relative to the dominant component to result in the underestimation of the decay lifetime; this is not observed on inspecting the lifetime estimates obtained on application of Bayesian decay analysis using the optimal IRF approximation, as shown in Fig. 8.15. The IRF parameter estimates offered by Bayesian SID analysis at lower intensities are, in this example, however, fairly poor as can be seen in Fig. 8.16. Although as intensity is increased the Bayesian SID algorithm becomes more likely to recognise the minor IRF component (true weighting of 0.2), even at an intensity of about 20000 total photon counts the Bayesian SID algorithm is most likely to capture only a single component IRF from the decay data. However, with a data set containing around 10^7 total photon counts, the optimal IRF approximation as determined by the Bayesian SID analysis of is in very close agreement with actual synthetic IRF, as shown in Fig. 8.14.

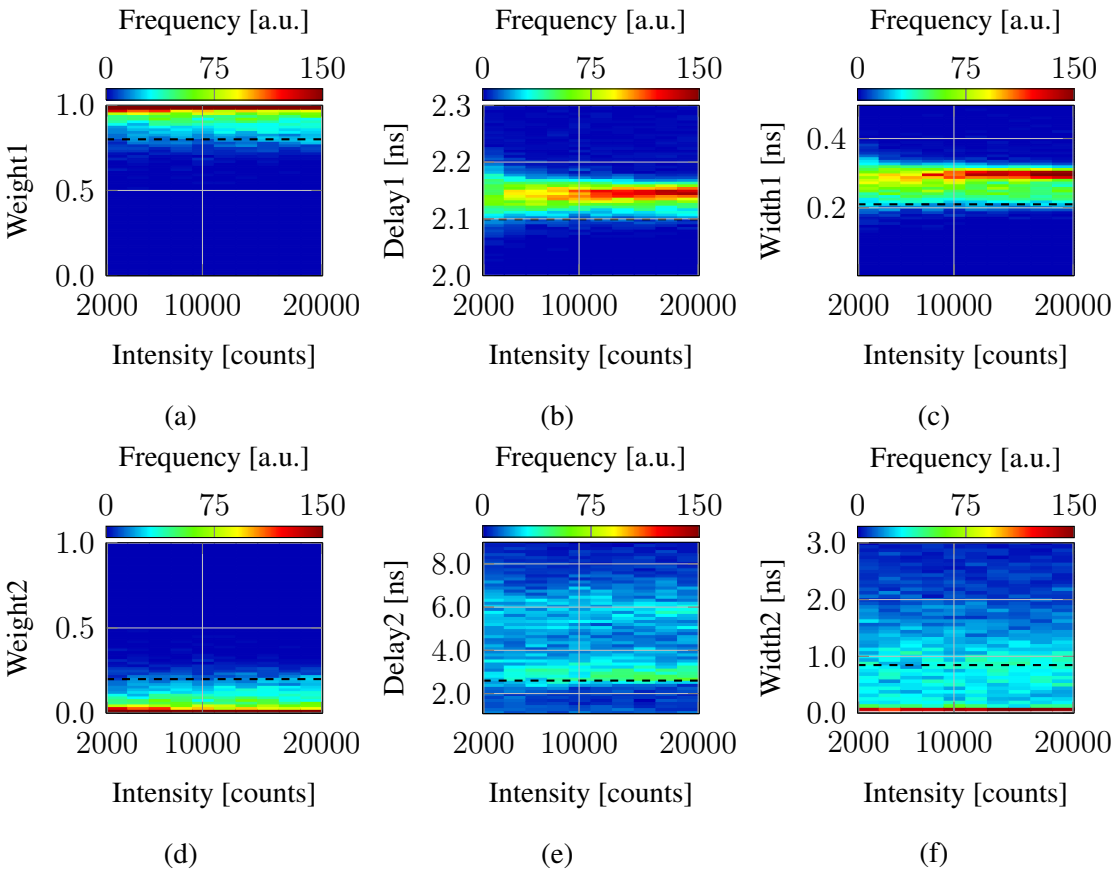


Figure 8.16: *Bayesian estimation of IRF parameters:* The Bayesian SID estimates of the two component IRF approximation parameters assuming a mono-exponential decay. In (a), (b), and (c), the parameter estimates that define the dominant component of the synthetic IRF, and in (d), (e), and (f), the parameter estimates that define the minor component.

8.3 Low count data from human epithelial carcinoma cells

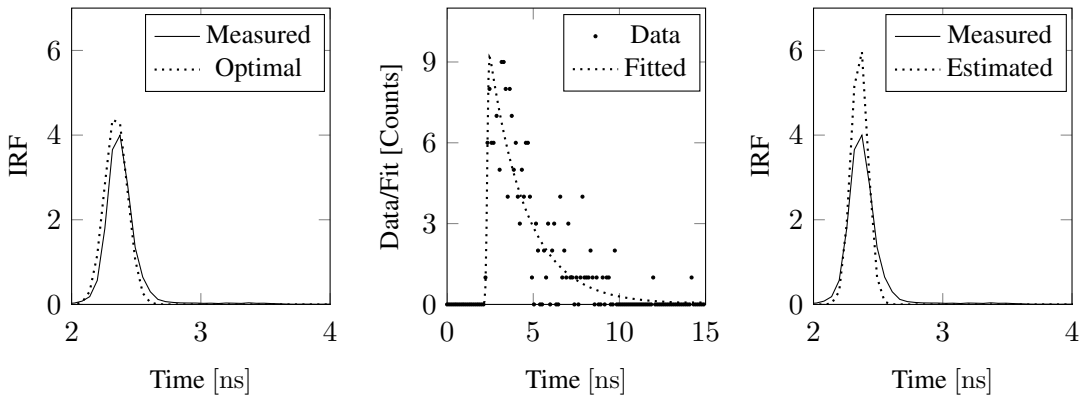
In this section the performance of the Bayesian SID algorithm is investigated using the low count mono-exponential data obtained for [1]. In studying the performance of Bayesian SID analysis on such low count experimental biological data both the obtained IRF parameter estimates and the lifetime estimates are examined. The IRF parameter estimates can be compared to the optimal[2] Bayesian single component IRF approximation obtained using a very high count data set (i.e. by binning all pixels in an image to yield a single time-resolved data set). The lifetime estimates are compared to those obtained from mono-exponential decay analysis using the optimal Bayesian single component IRF and also from ML; the 'true' lifetime is taken to be 2.18 ns, the value that all techniques were found to converge to in the analyses performed for [1].

The data were purposefully collected to have low intensities (relative to what would usually be expected for an analysis to be deemed trustworthy using standard techniques) at each image pixel; the data analysed for this section having between around 100 and 1000 total counts. Being mindful of the meagre total counts available, the results presented here are for Bayesian SID analysis under the assumption that the decay is purely mono-exponential and that the IRF can, at least to a reasonable first approximation, be represented by a single Gaussian. Although, of course, it may be that the real IRF would be more accurately represented by two or more Gaussian components, it is necessary to have realistic expectations as to what can be reliably[3] estimated given the short supply of data.

The optimal Bayesian single Gaussian IRF approximation, estimated from a data set of more than 10^7 photon counts, is shown in Fig. 8.17, having a width of 0.087 ns (FWHM 0.205 ns) and being centered around a delay of 2.343 ns. Despite even the optimal single

[2]Of course, it is quite probable that the optimal Bayesian single component IRF approximation is less a "gold" standard than a "silver" or even "bronze" standard, with more accurate approximations being realised with the introduction of additional Gaussian components; however, for ease of comparison with the estimates here the optimal single component IRF is employed as the ideal reference.

[3]It is sensible always to be vigilant against over-fitting and to balance what is desired from an analysis with what is possible given the data available. Although additional parameters are easily introduced to, say, model a bi-exponential decay and a multi-component IRF approximation, experience suggests that meaningful parameter estimates would be harder to come by from the analysis of such low photon count data using such a model.



(a) Bayesian optimal IRF (b) SID, typical pixel (c) IRF, typical pixel

Figure 8.17: *Bayesian SID analysis of high and low photon count data:* In (a) the experimentally measured IRF (normalised) and the optimal single Gaussian component Bayesian SID approximated IRF, estimated from a very high count data set (10^7 total counts); the single Gaussian being of width 0.087 ns (FWHM of 0.205 ns), centered around a delay of 2.343 ns. In (b) the data and Bayesian SID fitted decay curve at a pixel having a total of 366 photon counts. The mono-exponential lifetime estimated to be 2.14 ± 0.13 ns. In (c) the single Gaussian Bayesian estimated IRF approximation having a width of 0.064 ns (FWHM 0.151 ns) centered about a delay of 2.355 ns.

Gaussian approximation not showing a perfect correspondence with the measured IRF it certainly does not seem unreasonable to use such an approximation in this case, given that the measured IRF is (at least on visual inspection) very nearly symmetrical and not greatly different in appearance to a Gaussian, particularly given the anticipated uncertainty in parameter estimates due to the low count nature of the data.

The time resolved data from a typical pixel are shown in Fig. 8.17 along with the estimated decay and the estimated and measured IRF for single pixel data; on visual inspection of the time resolved data, it would seem reasonable to say that obtaining accurate parameter estimates from such data (having 366 total photon counts at the pixel chosen for this example) is likely to prove challenging. The Bayesian SID algorithm does, however, yield a reasonable approximation to the measured IRF and also yields a respectable lifetime estimate. The lifetime and IRF width and delay parameter estimates obtained using Bayesian SID analysis of image pixels having photon counts between about 350 counts and 3500 counts are shown in Fig. 8.18, with most of the analysed pixels having fewer than 1000 counts. It is clearly apparent that the Bayesian SID algorithm is able to

consistently extract reasonable IRF and decay parameter estimates from such low count data. The distribution of IRF delay estimates has an average of 2.342 ns and a standard deviation of 0.019 ns. The width estimates are distributed about an average of 0.086 ns with a standard deviation of 0.022 ns. The lifetime estimates are distributed about an average of 2.19 ns with standard deviation 0.08 ns.

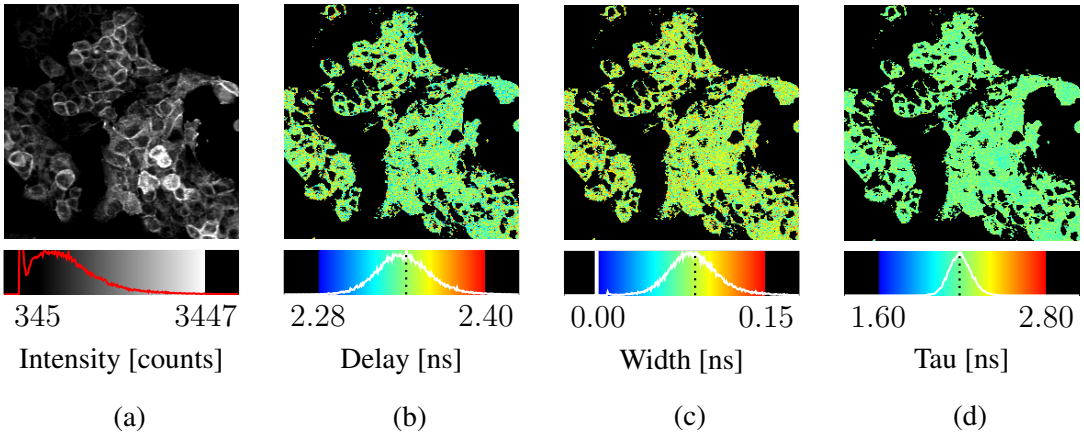


Figure 8.18: *Bayesian SID analysis of a low photon count image:* In (a) an intensity image of which pixels having a total photon count between about 345 and 3450 were analysed with Bayesian SID algorithm. In (b) and (c) the delay and width parameter estimates respectively that describe the Gaussian IRF approximation, and in (d) the mono-exponential decay lifetime estimates obtained using the Bayesian SID algorithm. The optimal Bayesian SID determined IRF delay and width parameter values and the 'true' decay lifetime are indicated by the dashed lines in the colour map beneath their respective images.

The reliability of the Bayesian SID algorithm IRF parameter estimates at low photon counts is studied in greater detail by collating a large number of estimates from the analysis of seven images. In Fig. 8.19 the IRF width and delay parameter estimates are shown for intensities between 100 and 1000 photon counts. It is clear that, even with the simplest mono-exponential decay and single component IRF approximation, the Bayesian SID algorithm struggles to estimate the IRF width parameter reliably below intensities of about 500 photon counts. The delay parameter, however, is reliably estimated even at intensities as low as 100 photon counts. It should also be observed that in those cases where Bayesian SID analysis significantly underestimates the IRF width, the delay parameter is also likely to be underestimated.

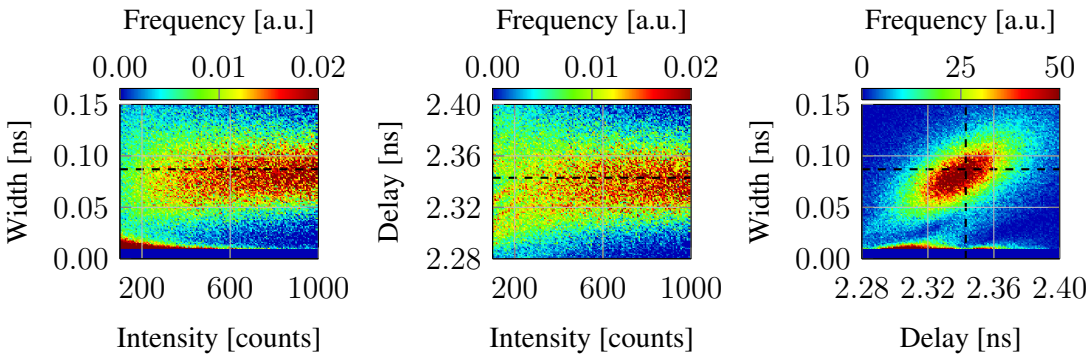
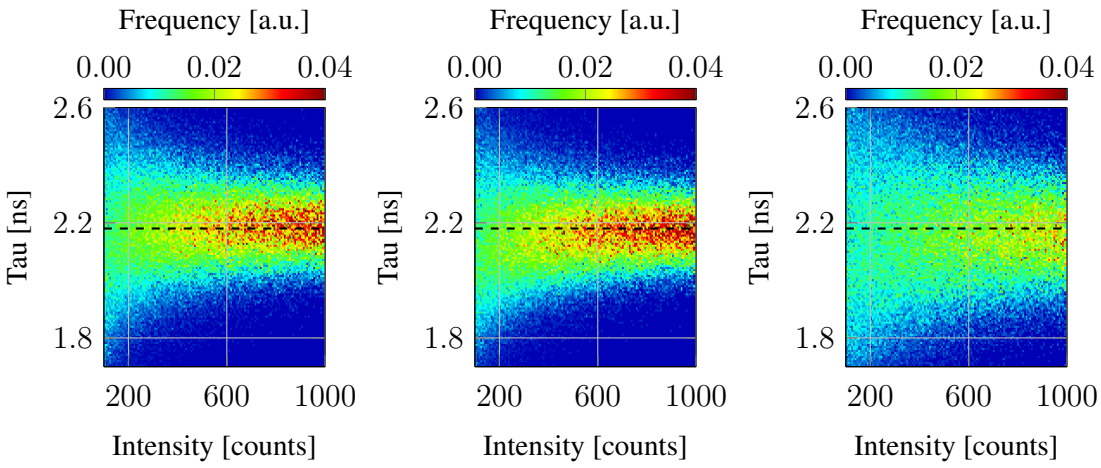


Figure 8.19: *Bayesian SID analysis IRF parameter estimation:* The width and delay parameter estimates that describe the single Gaussian IRF approximation employed for this mono-exponential Bayesian SID analysis (optimal IRF delay and width values are indicated by dashed lines). In (a) the estimated width of the Gaussian; observe that at below an intensity of about 500 total counts the Bayesian SID algorithm is far more likely to significantly underestimate the width, most often defaulting to the minimum permitted value in the implementation. At higher intensities the estimates are spread around the optimal value of 0.087 ns. In (b) the estimated delay parameter value around which the Gaussian is centered; although the estimates are most tightly distributed around the optimal value of 2.343 ns at the highest intensities (as would be expected), it should also be noted that there does not appear to be any systematic bias even at very low intensities. In (c) the interdependence of the delay and width parameter estimates; notice that when the width is significantly underestimated, the delay parameter also tends to be underestimated.

Considering now the mono-exponential decay lifetime estimates; it is pleasing to note that the lifetime estimates resulting from Bayesian SID analysis are not very different to those obtained from Bayesian decay analysis alone (using the optimal IRF approximation), as shown in Fig. 8.20. It is also striking that, despite the Bayesian SID lifetime estimates being very slightly degraded in comparison to those obtained from Bayesian decay analysis alone using the optimal approximated IRF, there remains a significant improvement over the estimates offered by ML (which uses the experimentally measured IRF). The lifetime images obtained using the Bayesian SID algorithm, Bayesian decay analysis alone with the optimal IRF approximation, and ML, are shown in Fig. 8.21 for an intensity image of which most of the analysed pixels having between about 70 and 250 photon counts; although the lifetime distributions of all three analysis types are centered

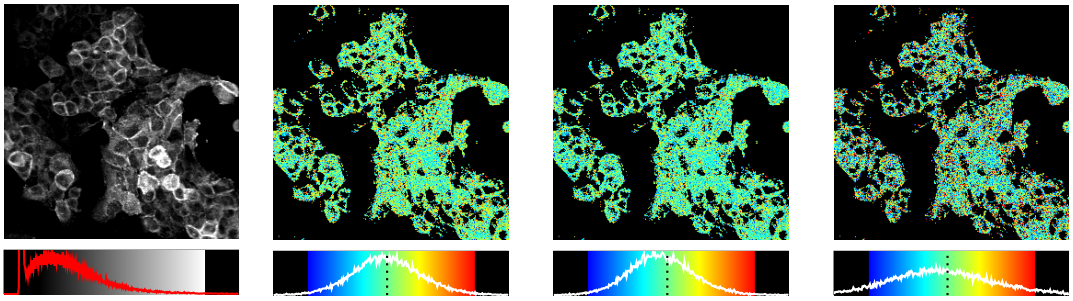


(a) Bayesian SID analysis (b) Bayesian decay analysis (c) ML decay analysis

Figure 8.20: *Comparing Bayesian SID analysis lifetime estimation (I):* In (a) the mono-exponential decay lifetime estimates obtained using Bayesian SID analysis under the assumption of a single component Gaussian IRF. In (b) the Bayesian lifetime estimates from mono-exponential decay analysis alone using the optimal single Gaussian IRF approximation parameter values. In (c) the lifetime estimates obtained using ML and the measured IRF. The 'true' decay lifetime of 2.18 ns is indicated by a dashed line in each plot, with the average and standard deviation of the lifetime distributions superimposed on the respective colour maps.

very closely around the true decay lifetime, the superiority of the Bayesian analyses (both SID and decay alone) over ML at low photon counts is clearly illustrated by the cleaner lifetime images. The Bayesian SID lifetime estimates distribution is centered about an average of 2.18 ns with standard deviation 0.17 ns. The estimates derived from Bayesian decay analysis alone (using the optimal IRF approximation) have an average of 2.16 ns and a standard deviation of 0.15 ns. The ML estimates are centered about an average lifetime of 2.17 ns with standard deviation 0.25 ns.

As demonstrated in Section 8.1, the IRF used in a decay analysis can have a significant influence on decay parameter estimates. It should therefore not be completely surprising that (unwanted) correlations may exist between Bayesian SID estimated IRF parameters and decay parameters. The relationships between the Bayesian SID estimated IRF parameters and the decay lifetime are illustrated in Fig. 8.22. It is evident (particularly on inspecting the histograms in (c) and (d) which are normalised across in order to reveal trends that are otherwise not so apparent) that underestimation of both the IRF delay and



(a) Intensity (b) Bayesian SID (c) Bayesian decay (d) ML decay

Figure 8.21: *Comparing Bayesian SID analysis lifetime estimation (II):* In (a) an intensity image of which pixels having an intensity between about 70 and 700 photon counts were analysed to yield the following mono-exponential decay lifetime images; in (b) the Bayesian SID algorithm decay lifetime estimates, in (c) the Bayesian decay analysis lifetime estimates (using the optimal Bayesian SID approximated IRF), and in (d) the ML lifetime estimates image (using the experimentally measured IRF). The 'true' decay lifetime of 2.18 ns is indicated in the estimated lifetime distribution for each analysis type.

width parameters is likely to be accompanied by an overestimated decay lifetime, whereas an overestimation of either IRF parameter is likely to be partnered by an underestimated decay lifetime.

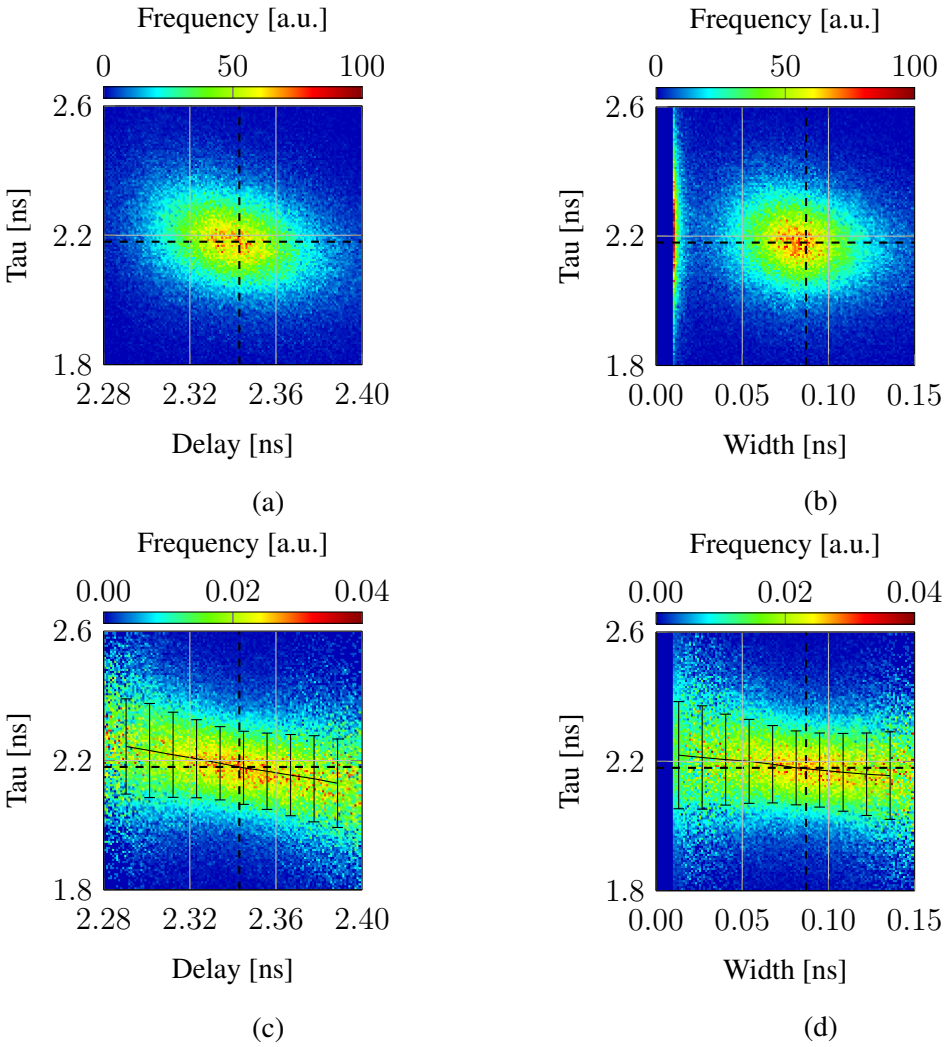


Figure 8.22: *Bayesian SID analysis, lifetime and IRF parameter correlations:* in (a) the relationship between the estimated lifetime and the delay parameter, and in (b) the relationship between the estimated IRF width and the decay lifetime. In (c) the estimated lifetime distribution conditioned on the IRF delay parameter, and in (d) the estimated lifetime distribution conditioned on the IRF delay parameter, showing that an overestimation of the decay lifetime is likely to be accompanied by an underestimation of the IRF width or/and decay parameter.

8.4 Advantages and weaknesses in using an analytic IRF model

It is clear that the IRF is important in the characterization of a decay signal from decay data and that it is wise to guard against the biasing of parameter estimates by using the best possible IRF available for analysis (whether measured or approximated). To what degree the IRF is important and to what extent parameter estimates are biased by the use of a wrong IRF in any analysis does, of course, depend on the specific characteristics of the data and the acquisition system as a whole (e.g. the IRF width, the bin width, etc.). It was shown in Section 8.1 that, even for a simple mono-exponential signal, using the wrong IRF in decay analysis can lead to poor estimation of decay parameters. The utility and effectiveness of Bayesian SID analysis, either for IRF determination or for decay analysis (where the real IRF is unknown and any estimates that realise its approximation are incidental), has been clearly demonstrated throughout this chapter using both synthetic IRFs and synthetic data and with real biological data acquired on a real system. In this section limitations and weaknesses inherent to Bayesian SID are considered, and where possible suggestions are made as to how these may be overcome.

As the Bayesian SID algorithm estimates the IRF and the decay parameters simultaneously it is important to consider to what extent unwanted correlations between IRF parameters and decay parameters could be present and how they may best be avoided. For example, will Bayesian SID analysis be able to distinguish between data due to a particular IRF width and lifetime, and the data due to a slightly faster decay but with a wider IRF? In order to reduce the impact of such issues it would seem prudent to incorporate a prior on the IRF parameters that is based upon coarse estimates derived from any measured IRF where available. It would be essential though, that in constructing any such prior that it not be so strong as to effectively remove the possibility of some possible IRF parameter estimates altogether. A possible compromise could be realised by a prior that exerts a fairly strong influence on the parameters defining the dominant mode of the IRF (as should be reasonably easily obtained by even a fairly crude analysis of a measured IRF), with the prior on the remaining parameters of other IRF components being relatively weak.

In developing the IRF approximation upon which the Bayesian SID analysis is founded (Section 4.4) the weighted superposition of any number of truncated Gaussian

distributions has been chosen as a suitable solution as, although it may not be possible to represent some arbitrarily jagged IRF[4], such an approach should be able to represent most real IRFs with a suitable number of Gaussian IRF components. The truncated Gaussian distribution is appealing as it offers the significant advantage that our solution (resulting from the convolution of the analytically approximated IRF with an exponential decay signal) remains fully analytic and incorporates error functions which ease the burden of implementation.

It is evident that a subtle flaw does exist in the analysis when considering the Bayesian SID algorithm at its most fundamental level; regardless of any apparent complexity, at the heart of the algorithm are simply a sum of exponentials (to describe the decay) and a sum of Gaussian distributions (to describe the IRF approximation). Although we can be fairly confident that a fluorescence decay signal does not contain any Gaussian-like component, the Bayesian SID analysis will falter should the IRF inherent in the data being analysed contain an exponential-like term as such would become associated with the decay signal itself.

Of course, theoretically, it may be that a real IRF can be perfectly well modelled by a superposition of (potentially many) such truncated Gaussian distributions. It is fairer to ask though, given that we are rarely likely to have an abundance of data, can a real IRF be adequately modelled by the superposition of just a *few* such distributions? The strength of the argument that it is possible to mimic almost any distribution with the superposition of enough Gaussian distributions diminishes as rapidly as the number of parameters required to describe those distributions grows. As elsewhere in this work, we therefore yearn for the simplest possible (IRF) model and appeal to Occam's Razor; instead enquiring as to how *few* Gaussian distributions are required to adequately describe a particular IRF?

[4]It should be noted that the desire for a fully analytic solution is driven not merely by a yearning for mathematical cleanliness, but in order to avoid the necessity to call upon computationally expensive numerical integration techniques in any implementation. Additionally, when using sufficient components to yield a good IRF approximation that adequately captures all pertinent features of the real IRF, it remains to be seen whether parameter estimates would be improved by the use of a more elaborate approximation other than in cases where the truncated Gaussian IRF approximation is obviously unsuitable, especially in the low photon count regime at which this work is primarily targeted.

Chapter 9

Bayesian decay model selection

In this chapter, the effectiveness of the Bayesian decay model selection to discriminate between mono- and bi-exponential decay data by *quantifying* the relative probability of each model, is demonstrated. Typically, a combination of factors will influence which decay model is chosen for the analysis of experimental data. In choosing the decay model the expectation of what it is *believed* to be is always likely to be significant, as is whether the intensity of the acquired experimental data is sufficient to support a statistically significant analysis using such a model. The selected model is often justified by inspecting the parameter estimates, and the closeness of the fits to the acquired data at a representative sample of image pixels.

The Bayesian decay model selection analysis, as developed in Chapter 5, can be applied to determine the model that most-likely underlies the analysed time-resolved data from any number of different models. In this work, a Bayesian decay model selection algorithm that distinguishes between mono- and bi-exponential decay data is presented, the Bayesian decay model selection algorithm having been developed to quantify the relative probabilities of the two models. The developed Bayesian decay model selection algorithm determines the most-probable decay model $\mathcal{H}^\star$ from an ensemble that includes only a mono-exponential decay model $\mathcal{H}_1$ and a bi-exponential decay model $\mathcal{H}_2$, as formalised below (from Chapter 4, Eqn. (4.6)),

- Mono-exponential decay model
  $\mathcal{H}_1 : \; p(s) = \theta(s) \, \frac{1}{\tau_1} \, e^{-s/\tau_1}, \quad \tau_1 > 0.$
- Bi-exponential decay model

$\mathcal{H}_2 : \; p(s) = \theta(s) \, \frac{1}{w_1+w_2} \left[\frac{w_1}{\tau_1} \, e^{-s/\tau_1} + \frac{w_2}{\tau_2} \, e^{-s/\tau_2} \right], \quad w_1 > 0, \; w_2 \geq 0, \; \tau_1 > 0, \; \tau_2 > 0.$

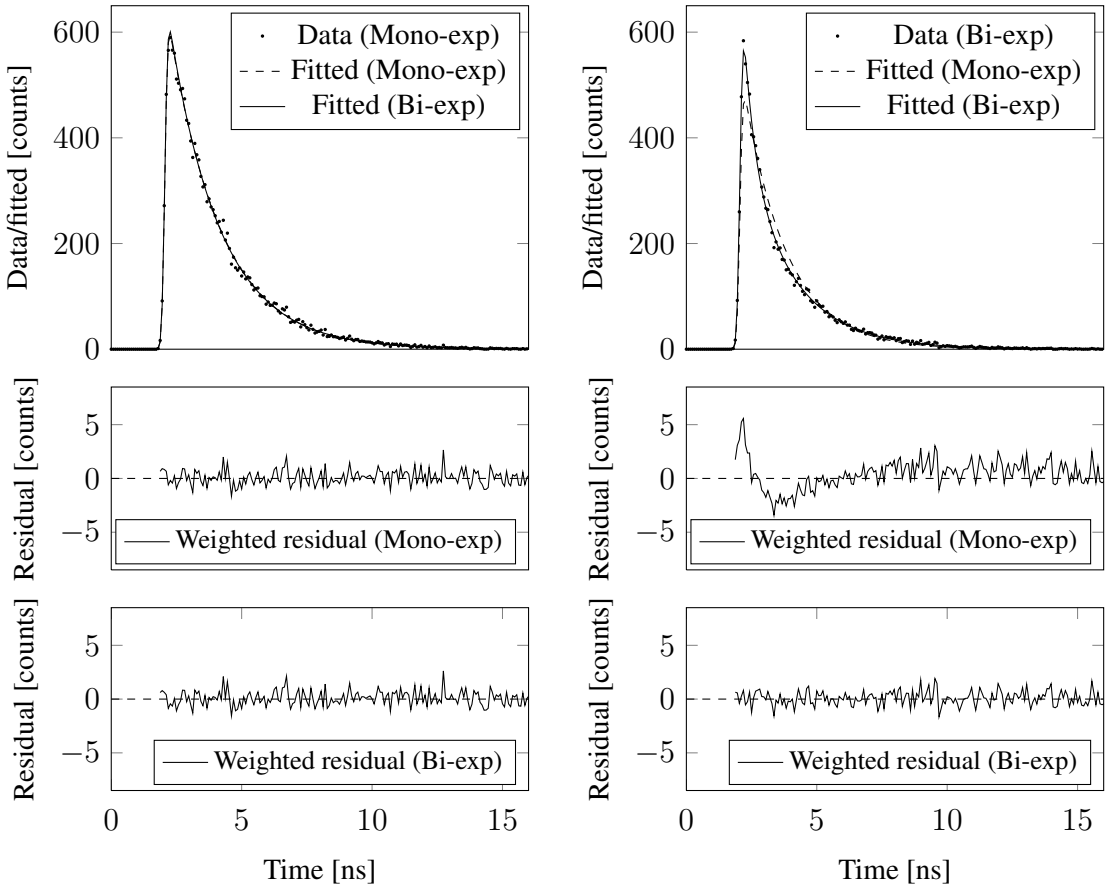
It is important to appreciate the need to be precise in the interpretation and description of the results of the developed Bayesian model selection algorithm; in this chapter the Bayesian model selection algorithm has been applied on the assumption that an ensemble including only mono-exponential and bi-exponential decay models is sufficient for the analysed data. Although not necessarily always explicitly stated, it should be understood that as the ensemble includes only mono-exponential and bi-exponential decay models, it is not possible to make inferences regarding the likelihood of, say, a tri-exponential decay (or some non-exponential decay process). Should there be a suspicion that decay data under consideration is due to, say, a tri-exponential decay then the Bayesian algorithm can be extended such that the relative likelihoods of a mono-exponential, bi-exponential, and tri-exponential decay model are determined.

The results of the application of Bayesian model selection to the experimental biological data, for which decay parameter analysis results were presented in Chapters 6 & 7, are presented in Section 9.2; Bayesian model selection thereby reinforcing the validity of the use of a mono-exponential decay model for the analysis of the human carcinoma cell data (Chapter 6) and a bi-exponential decay model for the analysis of the human breast cancer tissue data (Chapter 7). In Section 9.3 the performance of the Bayesian model selection algorithm at low intensities is considered, its effectiveness in correctly classifying time-resolved data as the FRET efficiency and interacting fraction of a synthetic bi-exponential decay are varied in Sections 9.3.1 & 9.3.2 respectively. However, before considering the performance of Bayesian decay model selection, the typical means of inferring which decay model is most appropriate is discussed in Section 9.1.

9.1 Decay model selection by visual inspection

In this section, the justification for using a particular decay model based on comparisons of the results of analysis using a selection of decay models is considered. The clues available for selecting the most appropriate decay model, or at least avoiding a clearly unsuitable decay model, from visual inspection of the analysis results are considered using two synthetically generated data sets, one simulating a mono-exponential decay and the

other a bi-exponential decay.



(a) Mono-exponential decay

(b) Bi-exponential decay

Figure 9.1: *Fitted decay and residuals as a visual aid for decay model selection:* In (a) synthetically generated mono-exponential decay data with accompanying fits and residuals from mono- and bi-exponential analysis, and in (b) the same for synthetic bi-exponential decay data. In (a) notice that it is not possible to resolve the mono- and bi-exponential fits to the synthetic mono-exponential data. Notice in (a) that there are no discernable differences between the residuals due to mono- and bi-exponential analysis of mono-exponential decay data. It is clear in (b) that the residuals from mono- and bi-exponential analysis of bi-exponential decay data differ considerably.

The fitted decays from a mono-exponential analysis and a bi-exponential analysis of the two data sets are shown in Fig. 9.1, along with their respective data-weighted residuals. The weighted residuals plot, which tracks the difference between the fitted and measured photon count at each bin, offers a convenient graphical method of inspecting how well

a particular model represents the data. Any obvious trends in the residuals distribution suggests an inappropriate model, whereas a model can be considered to appropriately fit the analysed data if the residual values are randomly distributed about the zero line.

There is no discernable difference between the fitted decays and their weighted residuals due to mono- and bi-exponential analysis of the mono-exponential decay data, as can be seen in Fig. 9.1 (a). Of course, a bi-exponential analysis is likely to yield a fitted decay that more closely corresponds to the data than a mono-exponential analysis is able to offer; the more parameters that a model has the easier it should be for it to closely fit some given data. However, differences between a mono-exponential fit and the bi-exponential model could be a consequence of the bi-exponential model "fitting to the noise" due to the additional flexibility it afforded by having more model parameters. In an effort to guard against such over-fitting it is usual to select the simplest of the models that adequately describes the data. On visual inspection of the fitted decays and their respective residuals, it would be likely that the data in Fig. 9.1 (a) would be correctly characterised as being due to a mono-exponential decay process.

As can be seen in Fig. 9.1 (b), inspection of the fitted decays due to mono-exponential and bi-exponential analysis of the bi-exponential decay data suggests that there is a clear difference between the two analyses. It is apparent that the mono-exponential fit does not follow the data as closely as does that due to bi-exponential analysis; the mono-exponential model has been unable to capture the effects of the fast decay component immediately following the peak of the transient. Mono-exponential analysis of the bi-exponential decay data yields an estimated mono-exponential decay lifetime of 1.65 ns, between the true bi-exponential decay lifetimes of 2.0 ns and 0.5 ns. It is obvious, on visual inspection, that the residual values due to bi-exponential analysis are randomly distributed about the zero line. Inspection of the residuals suggests that a mono-exponential model is not adequate for the description of the bi-exponential data; the mono-exponential residuals indicate that on the rise of the transient the fitted photon count is over-estimated, at the beginning of the decay of the fitted photon count is underestimated, and in the latter half of the measurement window the fitted photon count is consistently overestimated.

9.2 Applied to human breast cancer cells and tissue

In this section the effectiveness of the Bayesian decay model selection algorithm on its application to experimental biological data, the same data which were analysed for the decay parameter estimation results presented in Chapters 6 & 7 and assumed to be mono-exponential and bi-exponential respectively, is demonstrated.

The Bayesian decay model selection algorithm was configured to determine, at each image pixel, whether the time-resolved decay data were most likely to be due to a mono-exponential or a bi-exponential fluorescence decay process and to estimate the relative likelihood of the two decay models. The human epithelial *carcinoma* cells and the breast tissue sample were both imaged using the Galileo microscope, the time resolved data being acquired into 256 time bins (of equal width) that subdivide a measurement interval of 15 ns. As the intensity at each image pixel was insufficient for a reliable bi-exponential analysis to be performed, 7×7 spatial binning was invoked. The Bayesian determined optimal single Gaussian instrument response approximation was used in both analyses.

The Bayesian decay model selection analysis strongly echoes the expectation that the fluorescence emission from the human epithelial *carcinoma* cells expressing cdc42-GFP closely follows a mono-exponential decay (Fig. 9.2 (b,c)); 93% of the analysed image pixels being determined to be most likely due to a mono-exponential decay. On analysis of the 37440 analysed pixels, 34796 pixels (93%) were deemed to be be due to a mono-exponential decay model $\mathcal{H}_1$ and the remaining 2644 pixels (7%) classified as being due to a bi-exponential decay process $\mathcal{H}_2$. On analysis of the human breast cancer tissue data, Bayesian decay model selection is resoundingly supportive of the data being due to a bi-exponential decay model (or, at least, not being due to a mono-exponential decay model). All image pixels are classified as having bi-exponential decay data, and all with very close to 100% probability. It should be stressed that the Bayesian model selection algorithm can only determine the most likely model from those models against which the time-resolved data is tested. In this case, only mono-exponential and bi-exponential decay models are considered, that is, it is assumed that the decay process is not tri-exponential or of some other form.

The fitted decays and residuals from mono-exponential and bi-exponential analyses of both the human carcinoma cell data and the breast cancer tissue data are shown in Fig. 9.3; similar arguments to those applied in Section 9.1 also suggest that the human carcinoma

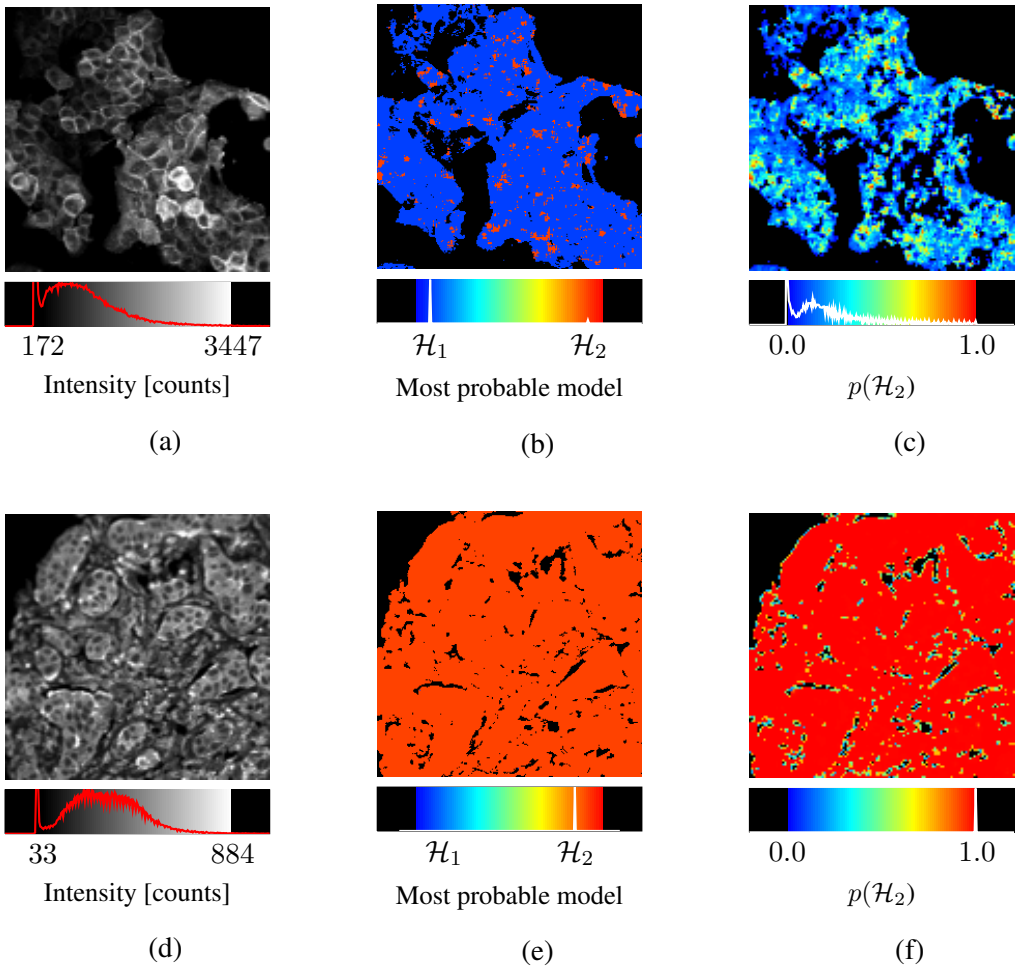
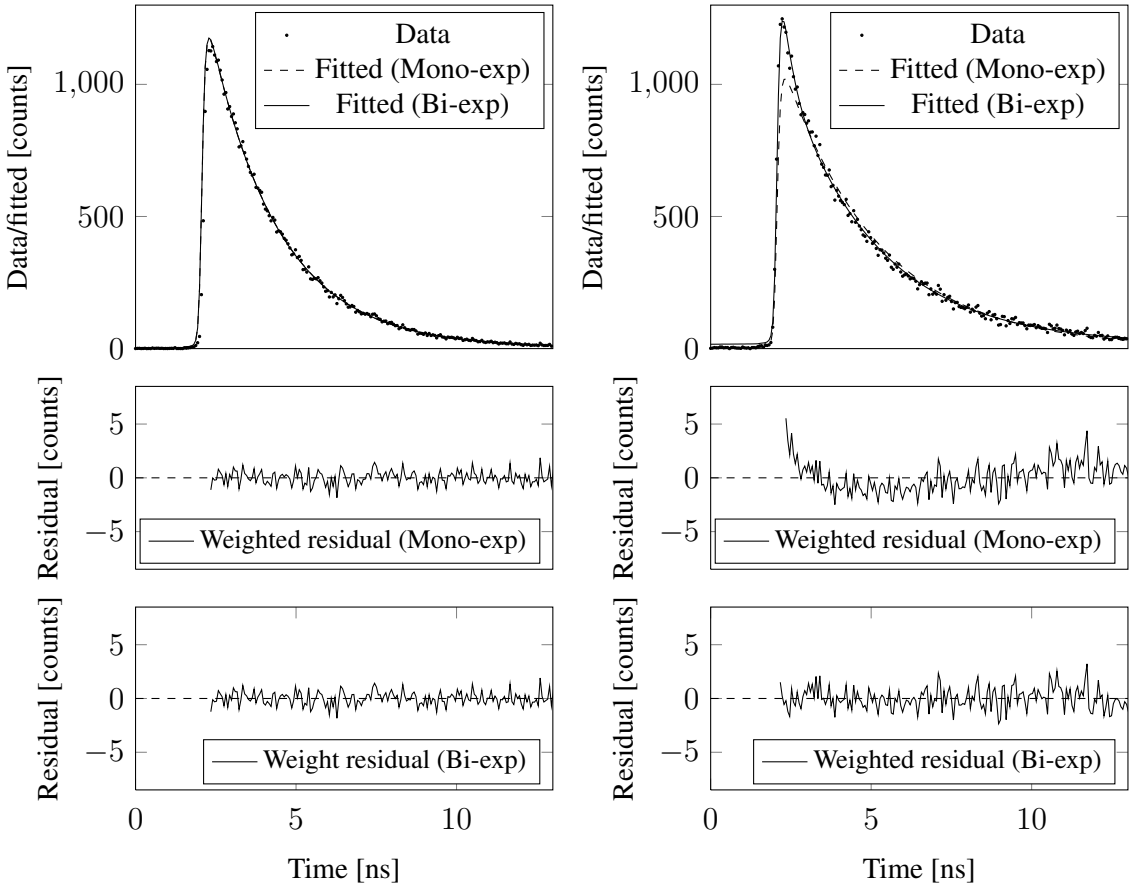


Figure 9.2: *Bayesian decay model selection applied to human breast cancer cell and tissue data:* In (a) an intensity image from the human epithelial *carcinoma* cell data analysed for Chapter 6 (mono-exponential analysis), in (b) the Bayesian determined most probable decay model map, and in (c) the relative probability of the decay being bi-exponential as opposed to mono-exponential. In (d) the intensity image of the human breast cancer tissue analysed for Chapter 7 (bi-exponential analysis), and in (e) and (f) the corresponding maps of the most probable decay model and the relative likelihood of a bi-exponential model as compared to a mono-exponential model respectively.

cell data are mono-exponential and the breast cancer tissue data are bi-exponential.



(a) Human carcinoma cell data

(b) Human breast cancer tissue data

Figure 9.3: *Fitted decays and residuals from mono-exponential and bi-exponential analysis of human breast cancer cell and tissue data:* Fitted decays and respective residuals from mono-exponential and bi-exponential analysis of the human carcinoma cell data in (a), and in (b) the same for the breast cancer tissue data.

9.3 Low count synthetic data

In this section the performance of the Bayesian model selection algorithm is assessed using synthetic data, for various FRET efficiencies in Section 9.3.1, for various interacting fractions in Section 9.3.2, and for binary FRET detection for very low photon count data in Section 9.4. The results presented in Sections 9.3.1 & 9.3.2 were obtained on application of the Bayesian model selection algorithm to the same synthetic bi-exponential data that had been subject to bi-exponential decay parameter analysis for Sections 7.1.1

& 7.1.2 in Chapter 7. In Section 9.4 the effectiveness of Bayesian decay model selection to characterise mono-exponential and bi-exponential decay data having intensities below about 1000 photon counts is demonstrated.

Unless stated otherwise, the Bayesian model selection algorithm was tested using data that simulated a TCSPC system having a repetition rate of 40 MHz and a measurement interval of 20.0 ns partitioned into 256 bins of equal width. All of the analysed transients were generated to incorporate the effects of a Gaussian instrument response, a uniform background of 0.5 counts per bin, and the addition of Poisson noise at each bin.

9.3.1 Performance and FRET efficiency

In this section the results of the application of the Bayesian model selection algorithm to synthetic bi-exponential data having various FRET efficiencies, at a fixed interacting fraction, are presented; the bi-exponential decay parameter analysis estimates for the same data was presented in Section 7.1.1. Transients with an intensity of about 10000 total photon counts and having different FRET efficiencies were generated by varying the ratio of the lifetimes of the bi-exponential decay components (with the lifetime of the slower decay component fixed at 2.0 ns), the initial amplitudes of the decay components being fixed to yield an interacting fraction of 0.5 in all cases.

The frequency with which the Bayesian decay model selection algorithm predicts a bi-exponential decay to be the most-probable model is shown in Fig. 9.4, along with the distribution of the probability assigned to the data being due to a bi-exponential decay. The Bayesian model selection algorithm correctly predicts a bi-exponential decay model for almost all of the analysed data when the FRET efficiency exceeds about 50%, that is, when the bi-exponential decay lifetime τ_1 of the slower decay component is more than double the lifetime τ_2 of the faster decay component. Bayesian model selection is more likely to predict the data to be mono-exponential when the ratio of the bi-exponential decay lifetimes are closer.

9.3.2 Performance and FRET interacting fraction

The results of the application of the Bayesian model selection algorithm to synthetic bi-exponential data having various interacting fractions, at a fixed FRET efficiency, are presented in this section; the bi-exponential decay parameter analysis estimates for the same

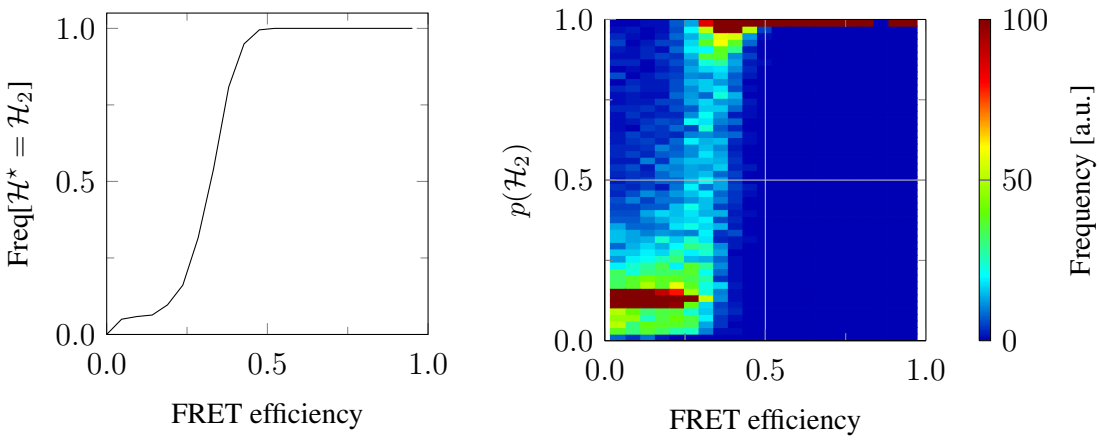


Figure 9.4: *Bayesian decay model selection and FRET efficiency:* In (a) the frequency with which the Bayesian model selection algorithm predicts a bi-exponential decay to be the most-probable model, and in (b) the distribution of the probability assigned to the bi-exponential model. Different FRET efficiencies were simulated by varying the ratio of the lifetimes of the bi-exponential decay components (the lifetime of the slower decay component being fixed at 2.0 ns), the initial amplitudes of the decay components being fixed to yield an interacting fraction of 0.5 in all cases. All of the analysed transients were generated to have an intensity of about 10000 total photon counts. At each different interacting fraction 2^{10} transients were analysed.

data was presented in Section 7.1.2. Transients having different interacting fractions were generated by varying the ratio of the initial amplitudes of the bi-exponential decay components, the lifetimes of the decay components were fixed at 2.0 ns and 0.5 ns yielding a FRET efficiency of 0.75. All of the analysed transients were generated to have an intensity of about 10000 total photon counts.

The frequency with which the Bayesian decay model selection algorithm predicts a bi-exponential decay to be the most-probable model is shown in Fig. 9.5, along with the distribution of the probability assigned to the data being due to a bi-exponential decay. It is evident that, at least at a FRET efficiency of 0.75 and an intensity of about 10000 total counts, the Bayesian decay model selection algorithm correctly predicts a bi-exponential decay model when the interacting fraction exceeds about 0.25. The Bayesian model selection algorithm does not correctly extract a bi-exponential model for interacting fraction values of less than about 20%, more frequently predicting instead that a mono-exponential

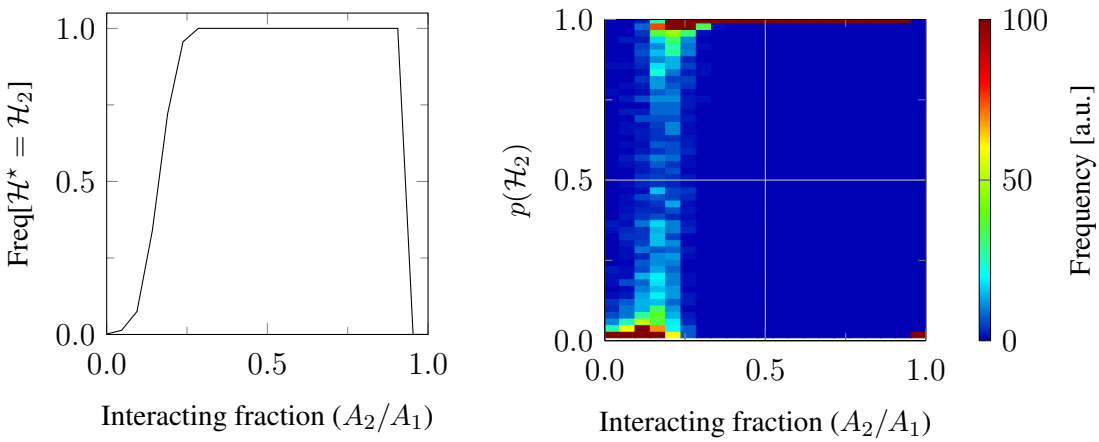


Figure 9.5: *Bayesian decay model selection and FRET interacting fraction:* In (a) the frequency with which the Bayesian model selection algorithm predicts a bi-exponential decay to be the most-probable model, and in (b) the distribution of the probability assigned to the bi-exponential model. Different interacting fractions were simulated by varying the ratio of the initial amplitudes of the two decay components, their lifetimes being fixed at 2.0 ns and 0.5 ns. All of the analysed transients were generated to have an intensity of about 10000 total photon counts. At each different interacting fraction 2^{10} transients were analysed.

decay is inherent in the data. It should also be noted that the Bayesian model selection algorithm correctly characterises the data as being due to a mono-exponential decay at the extremes of the tested interacting fraction values (i.e. 0% and 100%).

The failure of the Bayesian model selection algorithm to determine the decay model to be bi-exponential for interacting fractions smaller than about 0.3 is consistent with the broadening of the distributions of the bi-exponential decay parameter estimates as shown in Figs. 7.13 & 7.14 (Chapter 7). At interacting fractions lower than about 0.4 the uncertainty in the Bayesian estimated FRET efficiency starts to increase significantly, as shown in Fig. 7.15 (Chapter 7), and below interacting fractions of about 0.3 the accuracy of the Bayesian estimated FRET efficiency is poor; the Bayesian model selection algorithm prefers a bi-exponential model for description of the data but, as the larger uncertainty in the Bayesian estimated FRET efficiency and interacting fraction, not always the correct bi-exponential model.

9.4 Binary determination of the occurrence of FRET

In this section the potential for the use of Bayesian model selection to sort between low photon count mono-exponential and bi-exponential data is demonstrated. Mono-exponential transients were generated to have a decay lifetime of $\tau_1 = 2.0$ ns. Bi-exponential transients having a FRET efficiency of 75% and an interacting fraction of 0.5 were simulated by fixing the lifetimes of the decay components to be $\tau_1 = 2.0$ ns and $\tau_2 = 0.5$ ns and their initial amplitudes to be equal. All of the analysed transients simulated a background-free system. The same Bayesian determined optimal single Gaussian instrument response approximation was used for both the analysis of the mono-exponential and the bi-exponential data.

The performance of Bayesian model selection to classify the synthetic low count data as being mono-exponential or bi-exponential is shown in Fig. 9.6. The Bayesian model selection algorithm correctly identifies the mono-exponential decay data to be mono-exponential for more than 95% of analysed pixels, for intensities between about 100 and 1000 photon counts; there is a residual error rate of under 5%, which persists even as the intensity is increased, where the mono-exponential data are misclassified as being bi-exponential by the Bayesian model selection algorithm. The extraction of the correct decay model proves to be more challenging for bi-exponential decay data at such low intensities, however, at an intensity of about 500 photon counts the Bayesian model selection algorithm is able to correctly identify the decay data to be bi-exponential for more than 90% of analysed pixels. At intensities below about 200 photon counts the algorithm is more likely to predicted the bi-exponential decay data to be mono-exponential.

Inspection of typical mono-exponential and bi-exponential decay data at such a low intensity, and the corresponding fitted decays and their respective residuals, emphasises just how difficult a task it would be to determine by visual inspection that the data are actually due to a bi-exponential decay, as is illustrated in Fig. 9.7. The Bayesian model selection algorithm correctly classifies the mono-exponential decay data (which has an intensity of 478 photon counts) as being mono-exponential (Fig. 9.7 (a)), although it does assign an 11% probability that the data may be due to a bi-exponential decay; on inspecting the fitted mono-exponential and bi-exponential decays and their respective residuals, and applying the same reasoning as applied in Section 9.1, it would be almost impossible to disagree with the Bayesian prediction that the data is mono-exponential. The

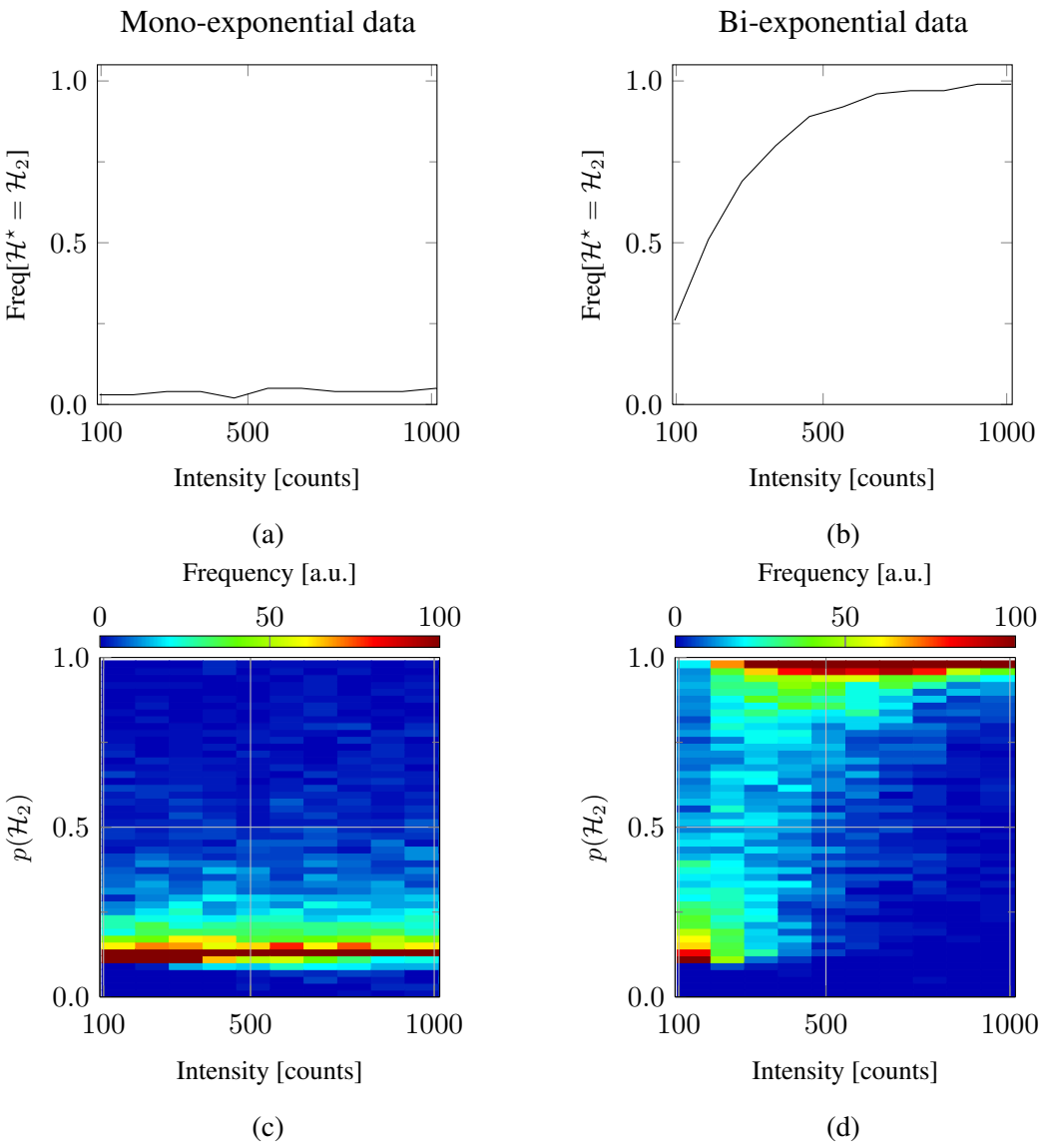
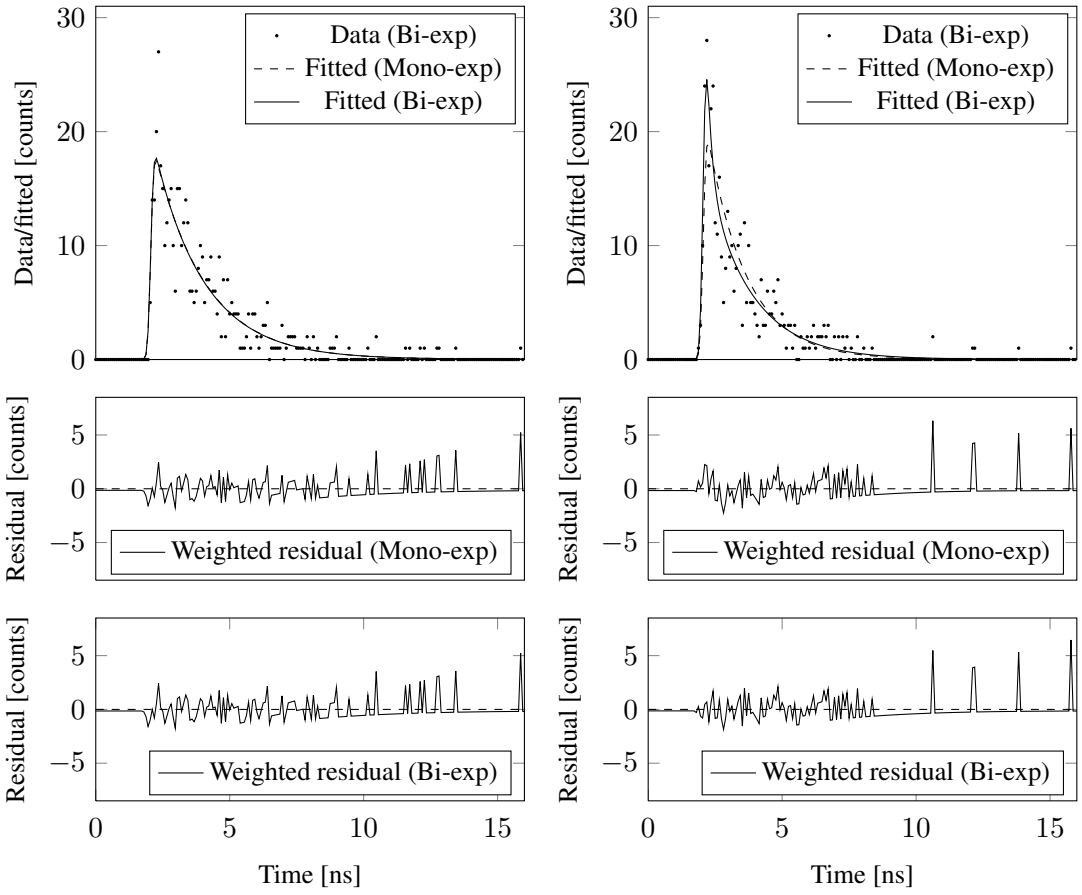


Figure 9.6: *Bayesian decay model selection for binary FRET determination:* In (a) and (b) the frequency with which the Bayesian model selection algorithm predicts a bi-exponential decay to be the most-probable model on application to to low intensity mono-exponential and bi-exponential data respectively, and in (c) and (d) the distribution of the probability assigned to the model being bi-exponential. Mono-exponential transients were generated to have a decay lifetime of $\tau_1 = 2.0$ ns. Bi-exponential transients having a FRET efficiency of 75% and an interacting fraction of one half were simulated by fixing the lifetimes of the decay components to be $\tau_1 = 2.0$ ns and $\tau_2 = 0.5$ ns and their initial amplitudes to be equal. At each intensity 2^{10} transients were analysed.



(a) Mono-exponential decay

(b) Bi-exponential decay

Figure 9.7: *Fitted decay and residuals as a visual aid for decay model selection at low intensities:* In (a) synthetically generated mono-exponential decay data with accompanying fits and residuals from mono-exponential and bi-exponential analysis, and in (b) the same for synthetic bi-exponential decay data. Notice that at an intensity of about 750 photon counts it is not possible to resolve the mono-exponential and bi-exponential fit to either the mono-exponential or bi-exponential synthetic data, nor are there any discernable differences between the residuals due to mono-exponential and bi-exponential analysis.

bi-exponential decay data has an intensity of 419 photon counts (Fig. 9.7 (b)); despite the transient appearing to be very challenging for any form of analysis, Bayesian model selection assigns a probability of 97% to the decay being bi-exponential and therefore correctly predicted the decay to be bi-exponential. Clearly, it would very difficult to confidently distinguish the decay as bi-exponential on the basis of visual inspection of the

fitted decays and their respective residuals; for these data the only comment that could be made in favour of a bi-exponential decay is that the bi-exponential residuals appear to be very slightly more randomly distributed about the zero line than do the mono-exponential residuals, the difference is though very slight.

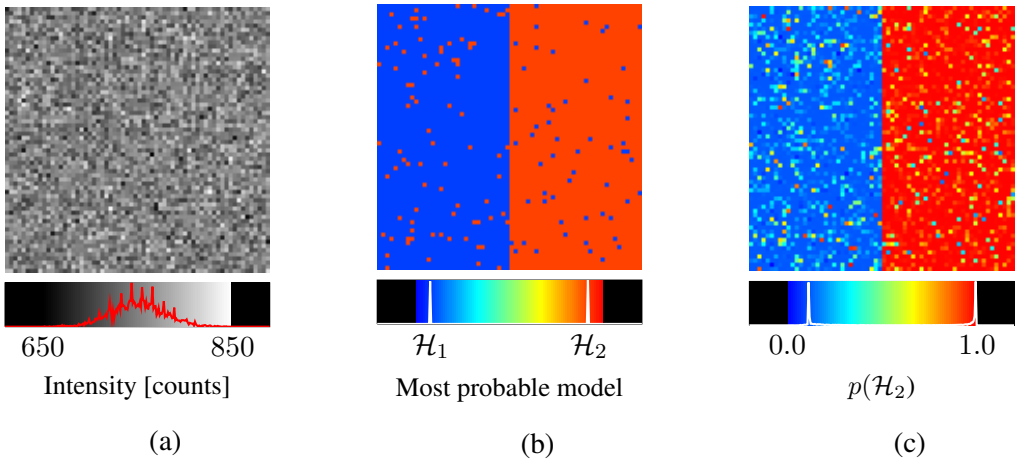


Figure 9.8: *Bayesian decay model selection for FRET detection:* Bayesian model selection images for the analysis of synthetic data having an intensity of about 750 photon counts. The pixels in the left half of the intensity image in (a) contain mono-exponential decay data, each generated to have a decay lifetime of $\tau_1 = 2.0$ ns, and the pixels in the right half contain bi-exponential decay data, each generated to have a FRET efficiency of 75% and an interacting fraction of one half by fixing the lifetimes of the decay components to be $\tau_1 = 2.0$ ns and $\tau_2 = 0.5$ ns and their initial amplitudes to be equal. In (b) and (c) the Bayesian predicted most probable decay model and the probability assigned to the data being bi-exponential respectively.

The Bayesian model selection algorithm is able to successfully distinguish between mono-exponential and bi-exponential decay data at very low intensities of about 750 photon counts, as shown in Fig. 9.8. The pixels in the left half of the image contain mono-exponential decay data, each generated to have a decay lifetime of $\tau_1 = 2.0$ ns, and the pixels in the right half contain bi-exponential decay data, each generated to have a FRET efficiency of 75% and an interacting fraction of 0.5 by fixing the lifetimes of the decay components to be $\tau_1 = 2.0$ ns and $\tau_2 = 0.5$ ns and their initial amplitudes to be equal. The analysed transients were generated to simulate a TCSPC system having a repetition rate of 40 MHz, a measurement interval of 20.0 ns partitioned into 256 bins of equal width, and incorporated the effects of a Gaussian instrument response, no background (i.e. uni-

form background of 0.0 counts per bin), and the effects of Poisson noise. Of the 2048 mono-exponential decays that were analysed 84 of them were incorrectly classified as being bi-exponential (i.e. false positives) and of the 2048 bi-exponential decays analysed 53 were classified as being mono-exponential (i.e. false negatives). Overall, the occurrence of FRET was predicted with about 97% accuracy by Bayesian decay model selection; an achievement that surely would not be possible by visual inspection of the fitted decays and their residuals, even should the FLIM data analyst be blessed with an abundance of time and even more patience.

Bayesian model selection could be used to sort between entities exhibiting either a mono-exponential or a bi-exponential decay where only low intensity data is available. Additionally, if the Bayesian decay model selection algorithm were to yield results sufficiently quickly it could form part of an adaptive imaging scheme where the dwell time at regions that are potentially undergoing FRET, having been classified as most-probably bi-exponential, is increased at the expense of regions that are unlikely to be undergoing FRET and have been classified as being most likely to be mono-exponential. Of course, for such an analysis to be effective the interacting fraction and FRET efficiency of the bi-exponential decay regions would need to be within the range where the model selection algorithm is known to work well. Such an analysis may prove to be potentially very useful in situations where the acquired data has insufficient photon counts for a bi-exponential decay analysis to yield parameter estimates with an uncertainty small enough for them to be useful, but where knowledge that FRET is occurring is beneficial.

Part IV

Discussion

The photon-by-photon Bayesian analysis for time-resolved FLIM data that has been presented in this thesis has been developed with the intention of providing the best possible analysis of TCSPC FLIM data. The desire for such a *gold standard* analysis may have initially been provoked by the need for more accurate mono- and bi-exponential decay parameter estimates with fewer photon counts, but as has been demonstrated the exploitation of the developed time-domain FLIM system model in the Bayesian framework is able to offer considerably more than just robust parameter estimation. In concluding this thesis, the developed Bayesian analysis for time-domain FLIM data is now reviewed, from its theoretical underpinnings through to its application to experimental biological data. The appeal of the developed fully-analytic time-domain FLIM system model is considered, and the results of its application in a Bayesian framework for the purposes of mono- and bi-exponential decay parameter estimation, instrument response determination, and decay model selection are summarised. The potential advantages and disadvantages that the developed Bayesian techniques may offer for FLIM data analysis are then discussed.

The fully analytic time-domain FLIM system model introduced in Part II incorporates a multi-exponential decay, an analytic approximation to the instrument response, and rigorously accounts for the effects of repetitive excitation that is inherent to a time-resolved system. The model describes the likelihood of a photon being counted into *any* of the time-bins within the entire measurement interval; the application of the FLIM system model for FLIM data analysis permits the use of *all* of the available data. The established direct-fitting analysis approaches that fit a decay to the measured photon count histogram sometimes operate only on the data between the peak of the transient and the end of the measurement window, information present in the data at the beginning of the measurement window sometimes being discarded and sometimes not being used correctly. In developing a Bayesian analysis for FLIM data one of the primary objectives was to reduce the number of photon counts required for a trustworthy analysis, thereby facilitating shorter imaging durations, it was therefore deemed essential that all available information should be used. Acquiring data having a sufficient intensity to yield useful parameter estimates is often difficult, and sometimes may not even be possible. It is obviously desirable to make the best possible use of all of those photon counts that have been acquired. In developing an analysis targeted at low intensity data, discarding useful information just because the analysis is unable to handle it is not an option.

In striving to obtain the greatest possible benefit from TSCPC acquired data, and

thereby the advantages the technique offers when performing FLIM to study the cellular environment, effective analysis is especially required when total photon counts are meagre. The benefit of the developed Bayesian analysis is most apparent at low total counts, potentially offering shorter imaging durations, or an improved spatial resolution by reducing the need for spatial binning, or both. In Part III the results of the application of the developed Bayesian analysis algorithms to low photon count data were presented; in Chapters 6 & 7 the developed Bayesian mono- and bi-exponential analysis algorithms were shown to consistently provide lifetime estimates with a greater accuracy than those of ML, LS, and phasor analysis, as discussed in greater detail below.

Bayesian mono-exponential analysis was shown to compare favourably with ML, LS, and phasor analysis in Chapter 6, as was initially reported in [1] for an analysis based on a FLIM system model that has since evolved. Bayesian mono-exponential analysis offers robust estimation with greater precision at very low total photon counts, particularly in the presence of significant background levels. Decay lifetime estimates within an accuracy of 20% with about 50 counts on the application of Bayesian mono-exponential analysis to synthetic decay data incorporating a 10% uniform background, this level of precision is not achieved with ML nor phasor analysis techniques with fewer than 100 photon counts. The analysis of low photon count data, obtained by imaging human epithelial carcinoma cells expressing cdc42-GFP, Bayesian analysis estimates the green fluorescent protein (GFP) lifetime to a level of accuracy not obtained using ML or other techniques. The robust performance and improved parameter estimates that Bayesian mono-exponential offers suggest that its use should always be seriously considered; rather than arguing for the use of Bayesian analysis on a case by case basis, such as for low count data, or for data including a significant background, perhaps it should be that if Bayesian analysis is not to be used for an analysis then a strong argument in favour of an alternative technique should be available.

The advantages offered by the developed implementation of Bayesian mono-exponential analysis are also highlighted in Chapter 6; the posterior distribution offers considerably more information than just the most likely decay parameter values. The Bayesian analysis implementation enables the posterior distribution to be explored graphically, and also offers the decay parameter average values and a reliable measure of the uncertainty associated with parameter estimates (Fig. 6.1). The visualisation of the posterior distribution provides an appealing means of assessing the reliability of decay param-

eter estimates, perhaps particularly appealing to a non-expert analyst, and could also offer a graphical means of determining when data having sufficient intensity had been acquired (Fig. 6.2). The uncertainty in the mono-exponential lifetime, as determined from the posterior distribution, was shown to correspond closely with the empirically-determined estimated uncertainty (Fig. 6.11); uncertainty maps indicate where within an image the most and least reliable lifetime estimates are situated, another visualisation that could prove useful both within a single image and also when applied to a sequence of images.

Bayesian bi-exponential analysis was shown in Chapter 7 to offer a modest improvement in the accuracy of the decay parameter estimates compared to those of ML and LS analysis. The Bayesian lifetime estimates were shown to converge to the true values at slightly lower intensities than did those of ML and LS analysis (Fig. 7.2), for synthetic data having a FRET efficiency of 75% and an interacting fraction of 0.5; for example, Bayesian estimates of the lifetime of the faster decay component were achieved within an accuracy of 20% at an intensity of about 5000 photon counts, estimates with such accuracy not being achieved until intensities of about 6500 and 7000 photon counts with ML and LS analysis respectively. The initial amplitude of the faster decay component also being determined with less uncertainty (Fig. 7.4). The greater accuracy in the bi-exponential decay parameter estimates yielding more accurate FRET efficiency and interacting fraction determination (Fig. 7.7) The Bayesian FRET efficiency and interacting fraction estimates also showed greater stability at very high FRET efficiencies than did those due to ML and LS analysis (Fig. 7.10). On the application of Bayesian bi-exponential analysis to breast cancer tissue data (Section 7.2) for comparison with ML and global ML analysis, a more general difficulty faced in analysing multi-exponential decay data was encountered; three different analysis methods each yielding very different results. In attempting to ascertain why the results of the different methods are not consistent with each other, inspection of the background estimates revealed only the Bayesian estimates to be consistent with the data (and expectation for the microscope system), and also suggest a spatial variation of FRET efficiency that ML and global ML analysis did not.

Although the results of the application of the Bayesian mono- and bi-exponential decay analysis algorithms demonstrate that they do offer a robust analysis for FLIM data with superior parameter estimates, it should be stressed that one of the key result of this work, perhaps as significant as the improved mono- and bi-exponential decay parameter estimates, has been the development of the fully analytic time-domain FLIM

system model. The developed fully analytic FLIM system model makes it possible to quantify how change in one of the FLIM system variables will influence the system as a whole. In this work, the fully analytic time-domain FLIM system model has been exploited for simultaneous fluorescence decay and instrument response parameter estimation, and also for fluorescence decay model selection without the need for numerical integration. Having already discussed the improved analysis offered by the Bayesian mono- and bi-exponential algorithms, attention now turns to the results of Bayesian SID analysis (Chapter 8) and Bayesian decay model selection (Chapter 9).

A Bayesian decay model selection algorithm has been applied to distinguish between mono- and bi-exponential decay data, quantifying the likelihood of the two models. On application to low intensity (about 750 photon counts) synthetic decays, Bayesian model selection was able to distinguish mono-exponential data from bi-exponential data with 97% accuracy, and may provide a means of determining where in a sample FRET is likely to be occurring even if the intensity is not sufficient to provide statistically relevant bi-exponential parameter estimates. The Bayesian model selection algorithm was demonstrated to function effectively on the experimental biological data that were analysed for Chapters 6 & 7; it is reassuring that 93% of analysed pixels of an image from the human carcinoma cell data (Chapter 6) were determined to be mono-exponential given that the samples were specifically prepared to yield a mono-exponential decay. On application to the breast cancer tissue data (Chapter 7), Bayesian model selection predicted that all of the analysed image pixels contain bi-exponential decay data, or at least that they do not contain mono-exponential decay data. The reliability and accuracy of bi-exponential parameter estimates were shown to depend heavily on the FRET efficiency and interacting fraction of the analysed decay (Chapter 7); Bayesian model selection was able to correctly classify bi-exponential data for decays having a FRET efficiency greater than about 50% and for decays having an interacting fraction of greater than about one third, values consistent with the limit of stability of bi-exponential parameter estimates (for all analysis methods, at least at the tested intensity), and could be used as a guide as to when a reliable bi-exponential analysis cannot be performed. As developed, the Bayesian decay model selection algorithm is only able to distinguish between mono- and bi-exponential decays; should it be desirable the implementation could be extended to offer model selection between mono-, bi-, and tri-exponential decays (or other decay models if really needed).

The simultaneous estimation of decay and instrument response parameters has been shown in Chapter 8 to be effective in determination of the instrument response approximation from the decay data alone. Indeed, for each of the different Bayesian algorithms, every result presented in Part III was obtained with an optimal instrument response approximation determined by the Bayesian SID algorithm. Although the Bayesian SID algorithm has been demonstrated to be useful both for determination of the optimal instrument response approximation and for decay analysis when the actual instrument response is not known, a drawback of the Bayesian SID algorithm as currently implemented is that it must be configured with the appropriate decay model and the number of Gaussian components that contribute to the instrument response approximation. A preferable approach, that uses the decay data, would be based on Bayesian model selection to determine the number of components required for the optimal instrument response approximation with the decay model being assumed. Alternatively, an even more ambitious approach that could be pursued for the determination of the optimal instrument response approximation would be to determine the most-probable decay model, the optimal number of Gaussian components and the most-probable parameter values for each component simultaneously.

The different Bayesian algorithms have been demonstrated to provide effective analysis, and in the case of decay parameter estimation Bayesian analysis has been shown to offer estimates that compare favourably with those from ML, LS, and phasor analysis. However, the results offered by Bayesian analysis do come at a cost; processing using the developed Bayesian analysis algorithms takes considerably longer than it does with the established methods, as summarised in Table 9.1.

Approximate analysis timings (per transient)				
Decay model	Bayesian	ML	LS	Phasor analysis
Mono-exponential	1.70 ms	0.17 ms	0.10 ms	0.03 ms
Bi-exponential	457.81 ms	0.51 ms	0.31 ms	-

Table 9.1: The approximate time required to perform Bayesian analysis, ML, LS, and phasor analysis. All results were obtained using a 2.53 GHz Intel Quad Core CPU with the C code compilation optimized for speed, average processing duration per transient is quoted from the analysis of synthetically generated data.

Clearly, Bayesian bi-exponential analysis of anything more than just a handful of such images would require the availability of considerable computing power or a lot of time

and patience, or most likely both. Putting the timings quoted in Table 9.1 into perspective for analysis of typical experimental data, it took about 64 seconds for a Bayesian mono-exponential analysis of a single image from the human carcinoma cell data, whereas ML, LS, and phasor analysis all provided estimates in under 16 seconds. The Bayesian bi-exponential analysis of the breast cancer tissue data (Section 7.2) took almost 8 hours; ML and LS analysis of the same data took about 37 seconds and 22 seconds respectively. Significant efforts have been directed towards reducing the analysis duration, and the figures quoted in Table 9.1 are, in fact, based on an optimised 'rapid' implementation of the Bayesian analysis algorithms. Given that it is when photon counts are in short supply that Bayesian analysis offers the greatest advantage, should photon counts be plentiful then it would be understandable if ML analysis would be preferred, waiting for Bayesian analysis with the advantage of only a slightly improved accuracy in the parameter estimates will appeal only to a very patient minority. Regardless of timing, though, the use of LS analysis could rarely be supported unless the data is of very high intensity, no alternative analysis implementations are available, and it is acknowledged that superior parameter estimates would be obtained with either ML or Bayesian analysis. Of course, as analysis of the breast cancer tissue data has demonstrated, sometimes it may be that the results of a Bayesian analysis may elucidate something that is not shown up by other methods, and sometimes the results may be worth the wait.

The Bayesian analysis implementation that was used to obtain the results presented here was developed primarily with robustness in mind. An accelerated Bayesian implementation could be realised by use of GPUs (graphical processing units) for faster computation, as has been achieved for number of medical imaging technologies [108], and recently for faster analysis in high-density localization microscopy [109]. Although in developing the Bayesian routines efforts were made to provide a 'rapid' implementation, it is certainly possible that the algorithms could be made considerably faster using optimised parallel computation. The current 'rapid' implementation of the Bayesian algorithms is, with only slight adaption, amenable to such an implementation.

Part V

Appendices

Appendix A

A fully analytic time-domain FLIM system model

The development of the time-domain FLIM system model, as introduced in Chapter 4, is presented here in greater detail. In Section A.1 a step-by-step development of the model is presented, intermediate steps that have not been presented in the main body of this work are outlined with a brief commentary as to the mathematical properties employed at each stage of the model development. In Section A.2 the development of Bayesian fluorescence decay model selection using a Gaussian approximation is given.

A.1 Step-by-step time-domain FLIM system model development

This section provides a step-by-step the development of the model presented in Chapter 4, offering a more expansive commentary of the model's development and detailing explicitly any mathematical tricks that have been employed.

A.1.1 Repetitive excitation in TCSPC FLIM

The photon arrival time data collected consists of the *recorded time* Δt on a periodic time window of duration T_m, of a photon with *emission time* t, that due to effects of the experimental apparatus, may have been delayed by a *delay time* u. Arrival times

are recorded only during a measurement interval of duration $T \leq T_m$, and therefore $\Delta t \in [0, T]$. Formalising this, the recorded photon arrival time determined with respect to the periodic window is given by,

$$\Delta t = t + u - T_m.\text{int}\left(\frac{t+u}{T_m}\right) \tag{A.1.1}$$

The probability of photon arrival time Δt being recorded in the measurement interval $[0, T]$ on the periodic window is given by,

$$p(\Delta t) = \theta(\Delta t)\theta(T - \Delta t)\left\{\frac{w_0}{T} + (1 - w_0)\frac{\int\int_0^\infty dtdu\, p(t)\Gamma(u)\delta\left(\Delta t - t - u + T_m.\text{int}\left(\frac{t+u}{T_m}\right)\right)}{\int_0^T d\Delta t' \int\int_0^\infty dtdu\, p(t)\Gamma(u)\delta\left(\Delta t' - t - u + T_m.\text{int}\left(\frac{t+u}{T_m}\right)\right)}\right\} \tag{A.1.2}$$

where the emission time t is distributed according to the decay signal $p(t)$, and the instrument response effects cause a delay u that is distributed according to $\Gamma(u)$ (it being assumed that the instrument can only introduce non-negative delays (i.e. $\Gamma(u) = 0$ if $u < 0$)). This expression can then be developed to the more generally useful expression of Eqn. (A.1.3), where the influence of recording arrival times with respect to the periodic window is captured by a summation. Rewriting the integral as follows,

$$\begin{aligned}
&\int_0^\infty du\, \Gamma(u)\, \delta\left(\Delta t - t - u + T_m.\text{int}\left(\frac{t+u}{T_m}\right)\right) \\
&= \sum_{\ell \geq 0} \int_0^\infty du\, \Gamma(u)\, \delta\left(\Delta t - t - u + \ell T_m\right) \delta_{\ell,\text{int}((t+u)/T_m)} \\
&= \sum_{\ell \geq 0} \Gamma(\Delta t - t + \ell T_m) \int_0^\infty du\, \delta\left(\Delta t - t - u + \ell T_m\right) \delta_{\ell,\text{int}((t+u)/T_m)} \\
&= \sum_{\ell \geq 0} \Gamma(\Delta t - t + \ell T_m)\theta[\Delta t - t + \ell T_m]\delta_{\ell,\text{int}((\Delta t+\ell T_m)/T_m)} \\
&= \sum_{\ell \geq 0} \Gamma(\Delta t - t + \ell T_m)\delta_{\ell,\text{int}(\ell+\Delta t/T_m)}\theta[\ell + (\Delta t - t)/T_m].
\end{aligned}$$

where $\delta(x)$ represents the Dirac delta function which exists only when x is equal to zero (i.e. $\int dx\ \delta(x)f(x) = f(0)$ for any function $f(x)$), $\delta_{i,j}$ is the Kronecker delta function which exists only when $i = j$, and the step function is denoted by $\theta(x)$ with $\theta(x > 0) = 1$ and $\theta(x \leq 0) = 0$. By definition, any observed photon arrival time $\Delta t \in [0, T]$ must fall within the measurement interval and therefore cannot exceed the modulation period T_m (i.e. $\Delta t \leq T \leq T_m$), consequently $\delta_{\ell,\mathrm{int}(\ell+\Delta t/T_m)} = \delta_{\ell,\ell} = 1$ and the integral simplifies further,

$$\begin{aligned}\int_0^\infty du\ \Gamma(u)\delta\left(\Delta t - t - u + T_m.\mathrm{int}\left(\frac{t+u}{T_m}\right)\right) \\ = \sum_{\ell\geq 0}\Gamma(\Delta t - t + \ell T_m)\theta[\Delta t - t + \ell T_m]\end{aligned}$$

Since the instrument response approximation $\Gamma(u)$ is non-zero only for non-negative delay times u (i.e. $\Gamma(u) = 0$ if $u < 0$), the above expression simplifies yet further, yielding,

$$\int_0^\infty du\ \Gamma(u)\delta\left(\Delta t - t - u + T_m.\mathrm{int}\left(\frac{t+u}{T_m}\right)\right) = \sum_{\ell\geq 0}\Gamma(\Delta t - t + \ell T_m)$$

Incorporating this result into Eqn. (A.1.2), the probability of photon arrival time Δt being recorded in the measurement interval $[0, T]$ on the periodic window is now given by,

$$p(\Delta t) = \theta(\Delta t)\theta(T - \Delta t)\left\{\frac{w_0}{T} + \frac{(1-w_0)\int_0^\infty dt\ p(t)\sum_{\ell\geq 0}\ \Gamma(\Delta t - t + \ell T_m)}{\int_0^T d\Delta t'\int_0^\infty dt\ p(t)\sum_{\ell\geq 0}\ \Gamma(\Delta t' - t + \ell T_m)}\right\}. \tag{A.1.3}$$

Denoting by $\Lambda(T, T_m)$ the normalisation constant in the term due to fluorescence decay photons, the photon arrival time likelihood (Eqn. (A.1.3)) can be written as,

$$p(\Delta t) = \theta(\Delta t)\theta(T-\Delta t)\left\{\frac{w_0}{T} + \frac{1-w_0}{\Lambda(T,T_m)}\int_0^\infty dt\; p(t)\sum_{\ell\geq 0}\Gamma(\Delta t - t + \ell T_m)\right\} \tag{A.1.4}$$

Notice that Eqn. (A.1.4) is completely general, incorporates a uniform background proportion w_0, the effects of repetitive excitation (through the summation), an arbitrary instrument response $\Gamma(u)$ and an arbitrary decay signal $p(t)$. The normalisation constant $\Lambda(T,T_m)$ is derived in Appendix A.1.5 for the general case (insofar as is possible) and then for the specific case where the instrument response and decay signal are of the form chosen in this work.

A.1.2 Discrete time nature of time-domain FLIM data

In accounting for the discrete time nature of the recorded arrival time in our model, the likelihood of photon arrival time Δt falling in an interval (i.e. a bin), with a bin being defined as the interval $b = [b^{\mathrm{L}}, b^{\mathrm{H}}] \subseteq [0,T]$, is required. Adopting the shorthand $p(b) = p(\Delta t \in b)$, the likelihood of a photon arrival time Δt falling in the bin $b = [b^{\mathrm{L}}, b^{\mathrm{H}}]$ is given by,

$$\begin{aligned}
p(b) &= \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t\; \theta(\Delta t)\theta(T-\Delta t)\left\{\frac{w_0}{T} + \frac{1-w_0}{\Lambda(T,T_m)}\int_0^\infty dt\; p(t)\sum_{\ell\geq 0}\Gamma(\Delta t - t + \ell T_m)\right\} \\
&= \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t\left\{\frac{w_0}{T} + \frac{1-w_0}{\Lambda(T,T_m)}\int_0^\infty dt\; p(t)\sum_{\ell\geq 0}\Gamma(\Delta t - t + \ell T_m)\right\} \\
&= \frac{w_0}{T}|b| + \frac{1-w_0}{\Lambda(T,T_m)}\int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t\int_0^\infty dt\; p(t)\sum_{\ell\geq 0}\Gamma(\Delta t - t + \ell T_m)
\end{aligned} \tag{A.1.5}$$

where the width of the interval is denoted by $|b| = b^{\mathrm{H}} - b^{\mathrm{L}}$. It is now convenient to define the general form of the "fluorescence decay bin-likelihood", $\hat{F}(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I})$, which describes the likelihood of a photon arrival time within the bin b due to the signal (i.e. not due to the uniform background),

$$\hat{F}(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) = \frac{1}{\Lambda(T, T_m)} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \int_0^{\infty} dt\, p(t) \sum_{\ell \geq 0} \Gamma(\Delta t - t + \ell T_m, \mathcal{I}) \tag{A.1.6}$$

with $\mathcal{I}$ denoting any parameters required as arguments due to the approximated instrument response. The bin-likelihood (Eqn. (A.1.5)) can now be written in the more compact form,

$$p(b) = \frac{w_0}{T}|b| + (1 - w_0)\hat{F}(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \tag{A.1.7}$$

The fluorescence decay bin-likelihood is developed for the chosen multi-exponential decay signal and approximation to the instrument response in Sections A.1.3 & A.1.4.

A.1.3 A multi-exponential decay signal

The introduction of a multi-exponential decay signal $p(t)$ of the form,

$$p(t) = \theta(t) \frac{\sum_{k=1}^{K} \frac{w_k}{\tau_k} e^{-t/\tau_k}}{\sum_{k=1}^{K} w_k}, \qquad w_k \geq 0 \; \forall k$$

into the general form of the fluorescence decay bin-likelihood (Eqn. (A.1.6)) yields the following,

$$\begin{aligned}
&\hat{F}(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \\
&\quad = \frac{1}{\sum_{k'=1}^{K} w_{k'}} \sum_{k=1}^{K} w_k \frac{1}{\Lambda(T, T_m)} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \int_0^{\infty} dt\, \tau_k^{-1} e^{-t/\tau_k} \sum_{\ell \geq 0} \Gamma(\Delta t - t + \ell T_m)
\end{aligned}$$

Since $p(t)$ is invariant under overall rescaling (i.e. $w_k \mapsto \mu w_k \; \forall \; k$), we may choose to require that $\sum_{k=1}^{K} w_k = 1 - w_0$ in order that the above can be rewritten as,

$$
\begin{aligned}
&\hat{F}(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \\
&\quad = \frac{1}{1-w_0} \sum_{k=1}^{K} w_k \frac{1}{\Lambda(T, T_m)} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \int_0^{\infty} dt\, \tau_k^{-1} e^{-t/\tau_k} \sum_{\ell \geq 0} \Gamma(\Delta t - t + \ell T_m) \\
&\quad = \frac{1}{1-w_0} \sum_{k=1}^{K} w_k F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I})
\end{aligned}
$$

where the quantity $F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I})$ is defined as,

$$
F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) = \frac{1}{\Lambda(T, T_m)} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \int_0^{\infty} dt\, \tau_k^{-1} e^{-t/\tau_k} \sum_{\ell \geq 0} \Gamma(\Delta t - t + \ell T_m, \mathcal{I}) \tag{A.1.8}
$$

with $\mathcal{I}$ denoting any parameters required as arguments due to the approximated instrument response. The bin-likelihood can now be written as,

$$
p(b) = |b| \frac{w_0}{T} + \sum_{k=1}^{K} w_k F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \tag{A.1.9}
$$

The fluorescence decay bin-likelihood $F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I})$ is determined for the chosen approximation to the instrument response in Section A.1.4.

A.1.4 Instrument response effects

In this section an instrument response approximation, comprising a weighted sum of truncated Gaussian distributions, of the form,

$$\Gamma(u,\mathcal{I}) = \sum_{i=1}^{I} \gamma_i \frac{e^{-\frac{1}{2}(u-u_i)^2/\sigma_i^2}}{\sigma_i\sqrt{2\pi}} \frac{2\theta[u-\delta_i]}{1+\mathrm{erf}((u_i-\delta_i)/\sigma_i\sqrt{2})},$$

$$\delta_i, u_i, \sigma_i \geq 0 \; \forall i, \quad \gamma_i \in [0,1] \; \forall i, \quad \sum_{i=1}^{I} \gamma_i = 1 \tag{A.1.10}$$

is introduced into the model. The quantity $\tilde{\gamma_i}$ is defined, for compactness,

$$\tilde{\gamma_i} = \gamma_i \; \Big(1+\mathrm{erf}((u_i-\delta_i)/\sigma_i\sqrt{2})\Big)^{-1} \tag{A.1.11}$$

and the instrument response is written as follows,

$$\Gamma(u,\mathcal{I}) = \sum_{i=1}^{I} \tilde{\Gamma}_i(u,\mathcal{I}), \qquad \tilde{\Gamma}_i(u,\mathcal{I}) = \frac{\sqrt{2}}{\sqrt{\pi}} \frac{\tilde{\gamma_i}}{\sigma_i} \, e^{-\frac{1}{2}(u-u_i)^2/\sigma_i^2} \, \theta[u-\delta_i]. \tag{A.1.12}$$

Incorporating the instrument response, the fluorescence decay bin-likelihood (Eqn. (A.1.8)) is decomposed into the sum of contributions of each of the instrument response components,

$$F(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) = \sum_i \tilde{F}_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \tag{A.1.13}$$

where $\tilde{F}_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I})$ represents the contribution of the ith instrument response component to the overall fluorescence decay bin-likelihood, and is given by,

$$\tilde{F}_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) = \frac{\tau^{-1}}{\Lambda(T,T_m)} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \sum_{\ell \geq 0} \int_0^{\infty} dt \, e^{-t/\tau} \, \tilde{\Gamma}_i(\ell T_m + \Delta t - t) \tag{A.1.14}$$

Determining now the convolution of a component of the multi-exponential decay signal,

having a decay lifetime τ, and the ith component of the instrument response approximation as follows,

$$
\begin{aligned}
&\int_0^\infty dt\; e^{-t/\tau}\; \tilde{\Gamma}_i(\ell T_m + \Delta t - t) \\
&\quad = \frac{\sqrt{2}}{\sqrt{\pi}}\; \frac{\tilde{\gamma}_i}{\sigma_i}\; \int_0^\infty dt\; e^{-\frac{1}{2}(\ell T_m + \Delta t - t - u_i)^2/\sigma_i^2 - t/\tau}\; \theta[\ell T_m + \Delta t - t - \delta_i] \\
&\quad = \frac{\sqrt{2}}{\sqrt{\pi}}\; \frac{\tilde{\gamma}_i}{\sigma_i}\; \theta[\ell T_m + \Delta t - \delta_i]\; \int_0^{\ell T_m + \Delta t - \delta_i} dt\; e^{-\frac{1}{2}(\ell T_m + \Delta t - t - u_i)^2/\sigma_i^2 - t/\tau} \\
&\quad = \frac{\sqrt{2}}{\sqrt{\pi}}\; \tilde{\gamma}_i\; e^{u_i/\tau}\; \theta[\ell T_m + \Delta t - \delta_i]\; e^{-(\ell T_m + \Delta t)/\tau}\; \int_{(\delta_i - u_i)/\sigma_i}^{(\ell T_m + \Delta t - u_i)/\sigma_i} du\; e^{-\frac{1}{2}u^2 + \sigma_i u/\tau} \\
&\quad = \frac{\sqrt{2}}{\sqrt{\pi}}\; \tilde{\gamma}_i\; e^{u_i/\tau + \sigma_i^2/2\tau^2}\; \theta[\ell T_m + \Delta t - \delta_i]\; e^{-(\ell T_m + \Delta t)/\tau}\; \int_{(\delta_i - u_i)/\sigma_i}^{(\ell T_m + \Delta t - u_i)/\sigma_i} du\; e^{-\frac{1}{2}(u - \sigma_i/\tau)^2} \\
&\quad = \frac{2}{\sqrt{\pi}}\; \tilde{\gamma}_i\; e^{u_i/\tau + \sigma_i^2/2\tau^2}\; \theta[\ell T_m + \Delta t - \delta_i]\; e^{-(\ell T_m + \Delta t)/\tau}\; \int_{(\delta_i - u_i)/\sigma_i\sqrt{2} - \sigma_i/\tau\sqrt{2}}^{(\ell T_m + \Delta t - u_i)/\sigma_i\sqrt{2} - \sigma_i/\tau\sqrt{2}} dv\; e^{-v^2} \\
&\quad = \tilde{\gamma}_i\; e^{u_i/\tau + \sigma_i^2/2\tau^2}\; \theta[\ell T_m + \Delta t - \delta_i]\; \tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i) \qquad \text{(A.1.15)}
\end{aligned}
$$

where the quantity $\tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i)$ is given by,

$$
\begin{aligned}
&\tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i) \\
&\quad = e^{-(\ell T_m + \Delta t)/\tau}\; \left[\mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right]
\end{aligned}
$$

Notice that the term $\theta[\ell T_m + \Delta t - \delta_i]$ ensures that the integral is positive. The developed expression describes (without normalisation) the likelihood of a fluorescence decay photon at time Δt in the measurement interval. The fluorescence decay likelihood (Eqn. (A.1.14)) due to the ith instrument response component can now be written as,

$$\tilde{F}_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I})$$
$$= \frac{\tau^{-1}}{\Lambda(T, T_m)} \tilde{\gamma}_i \, e^{u_i/\tau + \sigma_i^2/2\tau^2} \sum_{\ell \geq 0} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \; \tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i) \; \theta[\ell T_m + \Delta t - \delta_i] \tag{A.1.16}$$

Observe that the term $\theta[\ell T_m + \Delta t - \delta_i]$ ensures that the fluorescence photon likelihood is zero until the decay time $\ell T_m + \Delta t$ exceeds the cutoff parameter δ_i. In determining the remaining integral, which accounts for the discrete time nature of TCSPC data, it is convenient to incorporate the bin boundaries directly into the developed expression, as follows,

$$\begin{aligned}
&\int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \; \tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i) \; \theta[\ell T_m + \Delta t - \delta_i] \\
&\quad = \int_0^{\infty} d\Delta t \; \tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i) \; \theta[\Delta t - (\delta_i - \ell T_m)] \theta[\Delta t - b^{\mathrm{L}}] \theta[b^{\mathrm{H}} - \Delta t] \\
&\quad = \begin{cases} 0, & \text{if } b^{\mathrm{L}} \leq b^{\mathrm{H}} \leq \delta_i - \ell T_m \\ \int_{\delta_i}^{b^{\mathrm{H}}} d\Delta t \; \tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i), & \text{if } b^{\mathrm{L}} \leq \delta_i - \ell T_m \leq b^{\mathrm{H}} \\ \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t \; \tilde{\chi}_i(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i), & \text{if } \delta_i - \ell T_m \leq b^{\mathrm{L}} \leq b^{\mathrm{H}}, \end{cases}
\end{aligned} \tag{A.1.17}$$

It is evident that if the time bin lies entirely before the cutoff there is no likelihood of a fluorescence decay photon being counted into it, the likelihood of photon being counted into a time bin which straddles the cutoff is determined by integrating between the cutoff and the upper bin boundary, and if the time bin is entirely beyond the cutoff then the likelihood of a photon arrival time in the bin is determined by integrating between the bin boundary values. These conditions are encapsulated in the following,

$$
\begin{aligned}
&\int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t\ \tilde{\chi}(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i)\theta[\ell T_m + \Delta t - \delta_i] \\
&\quad = \theta[\ell T_m + b^{\mathrm{L}} - \delta_i] \int_{b^{\mathrm{L}}}^{\delta_i} d\Delta t\ \tilde{\chi}(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i) \\
&\qquad + \theta[\ell T_m + b^{\mathrm{H}} - \delta_i] \int_{\delta_i}^{b^{\mathrm{H}}} d\Delta t\ \tilde{\chi}(\ell, \Delta t, \tau, T_m, \sigma_i, \delta_i, u_i) \qquad \text{(A.1.18)}
\end{aligned}
$$

Focussing now on the remaining integral:

$$
\begin{aligned}
&\int d\Delta t\ e^{-(\ell T_m + \Delta t)/\tau} \left[\mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right] \\
&\quad = -\tau\ e^{-(\ell T_m + \Delta t)/\tau} \left[\mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right] \\
&\qquad + \tau \int d\Delta t\ e^{-(\ell T_m + \Delta t)/\tau} \frac{d}{d\Delta t} \left\{ \mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right\} \\
&\quad = -\tau\ e^{-(\ell T_m + \Delta t)/\tau} \left[\mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right] \\
&\qquad - \tau \int d\Delta t\ e^{-(\ell T_m + \Delta t)/\tau} \frac{d}{d\Delta t} \mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \\
&\quad = -\tau\ e^{-(\ell T_m + \Delta t)/\tau} \left[\mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right] \\
&\qquad + \tau \frac{2}{\sigma_i \sqrt{\pi}} \int d\Delta t\ \exp\left(-(\ell T_m + \Delta t)/\tau - \left[(u_i - \ell T_m - \Delta t)/\sigma_i \sqrt{2} + \sigma_i/\tau\sqrt{2} \right]^2 \right) \\
&\quad = \tau\ e^{-(\ell T_m + \Delta t)/\tau} \left[\mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right] \\
&\qquad + \tau \frac{2}{\sigma_i \sqrt{\pi}} e^{-u_i/\tau - \sigma_i^2/2\tau^2} \int d\Delta t\ e^{-\frac{1}{2}(u_i - \ell T_m - \Delta t)^2/\sigma_i^2} \\
&\quad = \tau\ e^{-(\ell T_m + \Delta t)/\tau} \left[\mathrm{erf}\left(\frac{(u_i - \ell T_m - \Delta t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \mathrm{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right] \\
&\qquad - \tau\ e^{-\sigma_i^2/2\tau^2 - u_i/\tau}\ \mathrm{erf}\left(\frac{u_i - \ell T_m - \Delta t}{\sigma_i \sqrt{2}} \right)
\end{aligned}
$$

Defining the quantity,

$$
\begin{aligned}
&\chi(\ell, t, \tau, T_m, \sigma_i, \delta_i, u_i) \\
&= e^{-(\ell T_m + t - u_i)/\tau + \sigma_i^2/2\tau^2} \left[\operatorname{erf}\left(\frac{(u_i - \ell T_m - t)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) - \operatorname{erf}\left(\frac{(u_i - \delta_i)\tau + \sigma_i^2}{\sigma_i \tau \sqrt{2}} \right) \right] \\
&\quad - \operatorname{erf}\left(\frac{u_i - \ell T_m - t}{\sigma_i \sqrt{2}} \right)
\end{aligned}
$$

the integral can be written as follows (using Eqn. (A.1.18)), whilst also absorbing the factor $\tau^{-1}\ e^{u_i/\tau + \sigma_i^2/2\tau^2}$ from Eqn. (A.1.16),

$$
\begin{aligned}
&\tau^{-1}\ e^{u_i/\tau + \sigma_i^2/2\tau^2} \int_{b^{\mathrm{L}}}^{b^{\mathrm{H}}} d\Delta t\ \tilde{\chi}_i(\ell, t, \tau, T_m, \sigma_i, \delta_i, u_i) \theta[\ell T_m + \Delta t - \delta_i] \\
&\quad = \theta[\ell T_m + b^{\mathrm{L}} - \delta_i] \left\{ \chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, b^{\mathrm{L}}, \tau, T_m, \sigma_i, \delta_i, u_i) \right\} \\
&\qquad + \theta[\ell T_m + b^{\mathrm{H}} - \delta_i] \left\{ \chi(\ell, b^{\mathrm{H}}, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i) \right\}
\end{aligned}
\tag{A.1.19}
$$

Consequently, the fluorescence decay bin-likelihood due to the ith instrument response component can be written as follows,

$$
\tilde{F}_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) = \frac{1}{\Lambda(T, T_m)} \tilde{\gamma}_i \sum_{\ell \geq 0} \Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) \tag{A.1.20}
$$

where the quantity $\Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i)$ is given by the following,

$$
\begin{aligned}
&\Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) = \\
&\quad \theta[\ell T_m + b^{\mathrm{L}} - \delta_i] \left\{ \chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, b^{\mathrm{L}}, \tau, T_m, \sigma_i, \delta_i, u_i) \right\} \\
&\quad + \theta[\ell T_m + b^{\mathrm{H}} - \delta_i] \left\{ \chi(\ell, b^{\mathrm{H}}, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i) \right\}
\end{aligned}
\tag{A.1.21}
$$

Incorporating this into the model yields the following expression for the bin-likelihood (which includes the contribution of both background and a fluorescence decay signal),

$$p(b) = |b|\frac{w_0}{T} + \sum_{k=1}^{K} w_k F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \mathcal{I}) \tag{A.1.22}$$

where $\mathcal{I} = \{\gamma_i, u_i, \sigma_i, \delta_i | i = 1, \ldots, I\}$ summarizes the parameters required for the approximation of the instrument response.

A.1.5 Normalisation of the fluorescence decay photon likelihood

The normalising factor $\Lambda(T, T_m)$ for the fluorescence decay photon likelihood term of Eqn. (A.1.4) depends on the specific form of the instrument response $\Gamma(u)$ and the decay signal $p(t)$ chosen. However, it can be determined so far in the general case of arbitrary instrument response and the decay signal, as follows,

$$\begin{aligned}
\Lambda(T, T_m) &= \int_0^T d\Delta t' \int_0^\infty dt\; p(t) \sum_{\ell \geq 0} \Gamma(\Delta t' - t + \ell T_m) \\
&= \int_0^\infty dt\; p(t) \sum_{\ell \geq 0} \int_0^T d\Delta t'\; \Gamma(\Delta t' - t + \ell T_m) \\
&= \int_0^\infty dt\; p(t) \sum_{\ell \geq 0} \int_{\ell T_m}^{\ell T_m + T} ds\; \Gamma(s - t) \\
&= \int_0^\infty dt\; p(t) \sum_{\ell \geq 0} \left\{ \int_{\ell T_m}^{\ell T_m + T_m} ds\; \Gamma(s - t) - \int_{\ell T_m + T}^{\ell T_m + T_m} ds\; \Gamma(s - t) \right\} \\
&= \int_0^\infty dt\; p(t) \int_0^\infty ds\; \Gamma(s - t) - \int_0^\infty dt\; p(t) \sum_{\ell \geq 0} \int_T^{T_m} ds\; \Gamma(\ell T_m + s - t) \\
&= 1 - \int_0^\infty dt\; p(t) \sum_{\ell \geq 0} \int_T^{T_m} ds\; \Gamma(\ell T_m + s - t)
\end{aligned} \tag{A.1.23}$$

In reaching Eqn. (A.1.23) it is demanded that the distributions $p(t)$ and $\Gamma(u)$ are themselves both normalised. Observe that when the measurement interval spans the entire modulation period (i.e. $T = T_m$) the normalisation constant $\Lambda(T, T_m) = 1$ as the integral term (Eqn. (A.1.23)) vanishes. However, usually, this cannot be assumed and the integrals must be determined, as is done for the chosen decay signal and instrument re-

sponse below. Rewriting Eqn. (A.1.23) to establish contact with the fluorescence decay bin-likelihood expression developed above,

$$\begin{aligned}\Lambda(T,T_m) &= 1-\int_0^\infty dt\; p(t)\sum_{\ell\geq 0}\int_T^{T_m} ds\;\Gamma(\ell T_m+s-t)\\ &= 1-\int_T^{T_m} ds\;\sum_{\ell\geq 0}\;\int_0^\infty dt\; p(t)\;\Gamma(\ell T_m+s-t)\\ &= 1-\frac{1}{\sum_\ell w_\ell}\sum_k w_k\int_T^{T_m} ds\;\sum_{\ell\geq 0}\;\int_0^\infty dt\;\tau_k^{-1}e^{-t/\tau_k}\;\Gamma(\ell T_m+s-t)\end{aligned}$$

Recognising that an integral of the form,

$$G(T,T_m,\tau_k,\mathcal{I}) = \int_T^{T_m} ds\;\sum_{\ell\geq 0}\;\int_0^\infty dt\;\tau_k^{-1}e^{-t/\tau_k}\;\Gamma(\ell T_m+s-t), \qquad \text{(A.1.24)}$$

has been determined above (Sections A.1.3 & A.1.4), having appeared first in Eqn. (A.1.8). The quantity to be determined in this case differs from that in Eqn. (A.1.8) in that, of course, the normalising factor $\Lambda(T,T_m)$ is absent, the integration variable s here replaces the variable Δt, and that the integral is over the interval $[T,T_m]$ rather than $[b^{\mathrm{L}},b^{\mathrm{H}}]$. It is therefore possible, unsurprisingly given that the interval $[T,T_m]$ is just another time-bin within the measurement interval (although not one into which photons are counted), to use the results developed above (Section A.1.3 & A.1.4) in reaching an expression for the normalisation constant $\Lambda(T,T_m)$.

$$\Lambda(T,T_m) = 1-\frac{1}{\sum_\ell w_\ell}\sum_k w_k G(T,T_m,\tau_k,\mathcal{I})$$

where the quantity $G(T,T_m,\tau_k,\mathcal{I})$ is given by the following,

$$G(T, T_m, \tau_k, \mathcal{I}) = \sum_{i=1}^{I} G_i(T, T_m, \tau_k, \mathcal{I}),$$
$$G_i(T, T_m, \tau_k, \mathcal{I}) = \tilde{\gamma}_i \sum_{\ell \geq 0} \Psi_i(\tau, T, T_m, T, T_m, \mathcal{I})$$

having used Eqn. (A.1.20), and the quantity $\Psi_i(\tau, T, T_m, T, T_m, \mathcal{I})$ is as defined by Eqn. (A.1.21). The remainder of this section explores some assorted properties of the normalisation constant and also offers some simple approximations.

Assorted properties of the normalisation constant

1. Property (i) Maximal measurement interval

$$T = T_m : \ \Lambda(T, T_m) \to 1 - \int_0^\infty dt \, p(t) \sum_{\ell \geq 0} \int_{(\ell+1)T_m - t}^{(\ell+1)T_m - t} du \, \Gamma(u) = 1 \quad \text{(A.1.25)}$$

2. Property (ii) Vanishing measurement interval

$$\lim_{T \to 0} \Lambda(T, T_m) \to 1 - \int_0^\infty dt \, p(t) \sum_{\ell \geq 0} \int_{\ell T_m - t}^{(\ell+1)T_m - t} du \, \Gamma(u)$$
$$= 1 - \int_0^\infty dt \, p(t) \int_{-t}^{\infty} du \, \Gamma(u) = 0 \quad \text{(A.1.26)}$$

3. Property (iii) Infinite modulation period

$$T = \alpha T_m, T_m \to \infty : \ \Lambda(T, T_m) \to 1 - \int_0^\infty dt \, p(t) \sum_{\ell \geq 0} \int_{(\ell+\alpha)T_m - t}^{(\ell+1)T_m - t} du \, \Gamma(u) = 1 \quad \text{(A.1.27)}$$

Simple approximation of the normalisation constant

In order to ease the implementation of the fluorescence decay photon bin likelihood a simple approximation for the normalization constant $\Lambda(T, T_m)$ is desired. Intuitively, the quantity $\Lambda(T, T_m)$ compensates for photons that cannot be detected as the system is 'blind' in the interval $[T, T_m]$. Inevitably, the likelihood of such photons being present but not detected depends on the decay signal itself and therefore it would seem reasonable that the lifetime of the decay signal be somewhere captured in any approximation. Additionally, the extent to which the 'tail' of the decay signal persists into the 'blind' period will depend on the extent to which the instrument retards the decay signal; it seems reasonable to neglect the effects of the instrument response other than the introduction of a delay to the observed decay signal (i.e. $\Gamma(u) = \delta(u - u_c)$ with $u_c < T$), and therefore,

$$
\begin{aligned}
&\hat{\Lambda}(T, T_m) \\
&\quad = 1 - \int_0^\infty dt\ p(t) \sum_{\ell \geq 0} \int_{\ell T_m + T - t}^{\ell T_m + T_m - t} du\ \delta(u - u_c) \\
&\quad = 1 - \int_{-\infty}^\infty du\ \ \delta(u - u_c) \int_0^\infty dt\ p(t) \sum_{\ell \geq 0} \theta[u - \ell T_m - T + t]\theta[\ell T_m + T_m - t - u] \\
&\quad = 1 - \int_0^\infty dt\ p(t) \sum_{\ell \geq 0} \theta[u_c - \ell T_m - T + t]\theta[\ell T_m + T_m - t - u_c] \\
&\quad = 1 - \sum_{\ell \geq 0} \int_{\ell T_m + T - u_c}^{\ell T_m + T_m - u_c} dt\ p(t)
\end{aligned}
$$

Notice that the expression developed so far satisfies all of the required properties as stated above.

$$
\int_a^b dt\ p(t) = \frac{1}{\sum_k w_k} \sum_k w_k (e^{-a/\tau_k} - e^{-b/\tau_k})
$$

Inserting this into the expression developed so far yields the following simple approximation to the normalising factor $\Lambda(T, T_m)$, as follows,

$$
\begin{aligned}
\hat{\Lambda}(T, T_m) &= 1 - \frac{1}{\sum_k w_k} \sum_k w_k (e^{-(T-u_c)/\tau_k} - e^{-(T_m-u_c)/\tau_k}) \sum_{\ell \geq 0} e^{-\ell T_m/\tau_k} \\
&= 1 - \frac{1}{\sum_k w_k} \sum_k w_k \left(e^{-(T-u_c)/\tau_k} - e^{-(T_m-u_c)/\tau_k}\right) \sum_{\ell \geq 0} \left(e^{-T_m/\tau_k}\right)^{\ell} \\
&= 1 - \frac{1}{\sum_k w_k} \sum_k w_k \frac{e^{-(T-u_c)/\tau_k} - e^{-(T_m-u_c)/\tau_k}}{1 - e^{-T_m/\tau_k}} \\
&= \frac{1}{\sum_k w_k} \sum_k w_k \left\{1 - \frac{e^{-(T-u_c)/\tau_k} - e^{-(T_m-u_c)/\tau_k}}{1 - e^{-T_m/\tau_k}}\right\} \\
&= \frac{1}{\sum_k w_k} \sum_k w_k \frac{1 - e^{-T_m/\tau_k} - e^{-(T-u_c)/\tau_k} + e^{-(T_m-u_c)/\tau_k}}{1 - e^{-T_m/\tau_k}}
\end{aligned}
\tag{A.1.28}
$$

which effectively averages the quantity $(1 - e^{-T/\tau})/(1 - e^{-T_m/\tau})$ over all signal components. A yet further simplification results from instead computing $\hat{\Lambda}(T, T_m)$ using the averaged signal component parameters, yielding the following,

$$
\tilde{\Lambda}(T, T_m) = \frac{1 - e^{-T_m/\langle\tau\rangle} - e^{-(T-u_c)/\langle\tau\rangle} + e^{-(T_m-u_c)/\langle\tau\rangle}}{1 - e^{-T_m/\langle\tau\rangle}}, \qquad \langle\tau\rangle = \frac{1}{K} \sum_{k=1}^{K} \tau_k
\tag{A.1.29}
$$

A.2 Fluorescence decay model selection using a Gaussian approximation

The model selection formulae developed in Chapter 5 are developed here for the FLIM system model of Chapter 4 and the fluorescence decay parameter prior distribution chosen in this work. In seeking the most probable fluorescence decay model $\mathcal{H}_K^\star$ and its hyperparameters $\boldsymbol{\alpha}_K^\star$, as given by (Eqns. (5.4, 5.6)),

$$(\mathcal{H}_K^\star, \boldsymbol{\alpha}_K^\star) = \underset{\mathcal{H}_K}{\operatorname{argmax}} \left[p(\mathcal{H}_K) p(\boldsymbol{\alpha}_K^\star|\mathcal{H}_K) \int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K|\mathcal{H}_K, \boldsymbol{\alpha}_K^\star) p(D|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \right] \tag{A.2.1}$$

where the optimal hyperparameter(s) $\boldsymbol{\alpha}_K^\star$ are determined according to,

$$\boldsymbol{\alpha}_K^\star = \underset{\boldsymbol{\alpha}_K}{\operatorname{argmax}} \left[p(\boldsymbol{\alpha}_K|\mathcal{H}_K) \int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K|\mathcal{H}_K, \boldsymbol{\alpha}_K) p(D|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \right] \tag{A.2.2}$$

it is necessary to define an appropriate model prior $p(\mathcal{H}_K)$ and hyperprior $p(\boldsymbol{\alpha}_K|\mathcal{H}_K)$. The simplest approach to choosing $p(\boldsymbol{\alpha}|\mathcal{H}_K)$ is to consider a maximum entropy prior for the decay lifetime(s) $\boldsymbol{\tau}_K$ with a fixed $\langle \tau_k \rangle$ (i.e. independent of k) having only one hyperparameter α, as given by,

$$p(\boldsymbol{\tau}_K) = \prod_{k=1}^{K} \alpha e^{-\alpha \tau_k} \theta[\tau_k] \tag{A.2.3}$$

In this case, as there is no rationale for the single hyperparameter α to be dependent on K, the hyperprior is independent of the decay model (i.e. $p(\boldsymbol{\alpha}_K|\mathcal{H}_K) = p(\alpha)$) and therefore, updating Eqns. A.2.1, A.2.2 accordingly yields,

$$(\mathcal{H}_K^\star, \alpha_K^\star)$$
$$= \underset{\mathcal{H}_K}{\text{argmax}} \left[p(\mathcal{H}_K) p(\alpha^\star) \int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K | \mathcal{H}_K, \boldsymbol{\alpha}_K^\star) p(D | \mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \right] \tag{A.2.4}$$

where the optimal hyperparameter value $\alpha^\star$ is given by,

$$\alpha_K^\star = \underset{\alpha>0}{\text{argmax}} \left[p(\alpha) \int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K | \mathcal{H}_K, \alpha) p(D | \mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \right] \tag{A.2.5}$$

In choosing a sensible $p(\boldsymbol{\alpha})$, the relationship between α and the average decay lifetime $\overline{\tau}$ is considered,

$$p(\tau_k) = \theta[\tau_k] \alpha e^{-\alpha \tau_k}, \qquad \overline{\tau} = \int_0^\infty d\tau \; p(\tau) \tau = \alpha^{-1} \tag{A.2.6}$$

Consider that, say, $\overline{\tau}$ is distributed homogeneously between $\overline{\tau} = 0$ and $\overline{\tau} = \tau_{\max}$, then,

$$\begin{aligned} p(\alpha) &= \int_0^{\tau_{\max}} d\overline{\tau} \; \frac{1}{\tau_{\max}} \delta \left(\alpha - \frac{1}{\overline{\tau}} \right) \\ &= \int_0^{\tau_{\max}} d\overline{\tau} \; \frac{1}{\tau_{\max}} \frac{1}{|\overline{\tau}^{-2}|} \delta \left(\overline{\tau} - \frac{1}{\alpha} \right) \\ &= \frac{1}{\alpha^2} \cdot \frac{1}{\tau_{\max}} \int_0^{\tau_{\max}} d\overline{\tau} \; \delta \left(\overline{\tau} - \frac{1}{\alpha} \right) \\ &= \frac{1}{\alpha^2} \cdot \frac{1}{\tau_{\max}} \theta \left[\alpha - \frac{1}{\tau_{\max}} \right] \\ &= \frac{\alpha_{\min}}{\alpha^2} \theta \left[\alpha - \alpha_{\min} \right] \end{aligned} \tag{A.2.7}$$

Inserting this into Eqn. (A.2.5), the optimal hyperparameter value $\alpha^\star$ is given by,

$$\alpha_K^\star = \underset{\alpha>\alpha_{\min}}{\text{argmax}} \left[\frac{1}{\alpha^2} \int d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K | \mathcal{H}_K, \alpha) p(D | \mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \right] \tag{A.2.8}$$

A.2.1 Integral determination using a Gaussian approximation

In developing a Bayesian analysis for fluorescence decay model selection (Chapter 5), the difficulties inherent in the computation of the required integrals (Eqn. 5.3),

$$\begin{aligned} Z_K &= \int_{\Omega_K} d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K|\mathcal{H}_K, \alpha) p(D|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \\ &= \int_{\Omega_K} d\boldsymbol{w}_K d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K|\mathcal{H}_K, \alpha) \prod_{j=1}^{M} p(b_j|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I})^{c_j} \end{aligned} \tag{A.2.9}$$

have been eased by the use of a Gaussian approximation (see e.g. [110, 111]). This section details the derivation of the second derivatives (i.e. the Hessian) required in such an approximation. Defining $S(\boldsymbol{w}_K, \boldsymbol{\tau}_K) = -\ln\left[p(\boldsymbol{w}_K, \boldsymbol{\tau}_K) p(D|\boldsymbol{w}_K, \boldsymbol{\tau}_K)\right]$, the posterior distribution can be written as,

$$\begin{aligned} &p(\boldsymbol{w}_K, \boldsymbol{\tau}_K|D) = \\ &\frac{p(\boldsymbol{w}_K, \boldsymbol{\tau}_K) p(D|\boldsymbol{w}_K, \boldsymbol{\tau}_K)}{\int d\boldsymbol{w}_K \int d\boldsymbol{\tau}_K \; p(\boldsymbol{w}_K, \boldsymbol{\tau}_K) p(D|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} = \frac{1}{Z_K} \, e^{-S(\boldsymbol{w}_K, \boldsymbol{\tau}_K)}, \qquad \boldsymbol{w}, \boldsymbol{\tau} \in \Omega_K \end{aligned} \tag{A.2.10}$$

where Z_K is a normalization constant (termed the 'evidence'). In making the Gaussian approximation, a Taylor expansion is taken around the peak of the Gaussian, located at $(\boldsymbol{w}_K^\star, \boldsymbol{\tau}_K^\star)$, such that,

$$\begin{aligned} &S(\boldsymbol{w}_K, \boldsymbol{\tau}_K) \approx \\ &S(\boldsymbol{w}_K^\star, \boldsymbol{\tau}_K^\star) + \frac{1}{2}((\boldsymbol{w}_K, \boldsymbol{\tau}_K) - (\boldsymbol{w}_K^\star, \boldsymbol{\tau}_K^\star))^{\mathrm{T}} \mathrm{A} ((\boldsymbol{w}_K, \boldsymbol{\tau}_K) - (\boldsymbol{w}_K^\star, \boldsymbol{\tau}_K^\star)), \quad \boldsymbol{w}, \boldsymbol{\tau} \in \Omega_K \end{aligned} \tag{A.2.11}$$

where the Hessian matrix and linear terms are absent as the the expansion is taken at the minimum of $S(\boldsymbol{w}_K, \boldsymbol{\tau}_K)$. In applying such an approximation, the 'evidence' Z_K is approximated by the quantity $Z_K^\star$ as given by,

$$Z_K^\star = e^{-S(\boldsymbol{w}_K^\star, \boldsymbol{\tau}_K^\star)} (2\pi)^K (\det \mathrm{A})^{-1/2}. \tag{A.2.12}$$

Consequently, the task of computing an integral now involves finding the derivatives that form the matrix A, which in this work is performed analytically.

A.2.2 Determining the Hessian

The Hessian in the case of fluorescence decay model selection is the square matrix of second-order partial derivatives, with respect to signal component weights w_x and lifetime τ_y where $x, y \in \{1, \ldots, K\}$, of the minus log probability, as given by,

$$S(\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) = \alpha \sum_k \tau_k - \sum_j c_j \ln p(b_j | \mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \tag{A.2.13}$$

where the leading term is due to the incorporation of a maximum entropy prior on the fluorescence decay parameters, and the bin-likelihood $p(b_j | \mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I})$ is as defined above (Eqn. (A.1.22)). Defining the Hessian elements in such a way that elements containing weight only derivatives appear in the top left quadrant, lifetime only derivatives in the bottom right quadrant, and cross derivatives elsewhere,

$$A = \begin{pmatrix} A^{(w,w)} & A^{(w,\tau)} \\ A^{(\tau,w)} & A^{(\tau,\tau)} \end{pmatrix}, \qquad A_{\mu\nu} = \begin{cases} \frac{\partial^2 S}{\partial w_\mu \partial w_\nu}, & \text{if } \mu \leq K, \nu \leq K \\ \frac{\partial^2 S}{\partial \tau_\mu \partial w_\nu}, & \text{if } \mu \leq K, \nu > K \\ \frac{\partial^2 S}{\partial w_\mu \partial \tau_\nu}, & \text{if } \mu > K, \nu \leq K \\ \frac{\partial^2 S}{\partial \tau_\mu \partial \tau_\nu}, & \text{if } \mu > K, \nu > K \end{cases} \tag{A.2.14}$$

Clearly, $A^{(w,\tau)}$ is the transpose of $A^{(\tau,w)}$, and therefore the following three derivatives are required,

$$\frac{\partial^2 S}{\partial w_\mu \partial w_\nu} = - \sum_j c_j \left\{ \frac{\partial^2 p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial w_\mu \partial w_\nu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} - \frac{\partial p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial w_\mu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} \frac{\partial p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial w_\nu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} \right\} \tag{A.2.15}$$

$$\frac{\partial^2 S}{\partial \tau_\mu \partial w_\nu} = - \sum_j c_j \left\{ \frac{\partial^2 p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial \tau_\mu \partial w_\nu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} - \frac{\partial p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial \tau_\mu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} \frac{\partial p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial w_\nu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} \right\} \tag{A.2.16}$$

$$\frac{\partial^2 S}{\partial \tau_\mu \partial \tau_\nu} = - \sum_j c_j \left\{ \frac{\partial^2 p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial \tau_\mu \partial \tau_\nu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} - \frac{\partial p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial \tau_\mu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} \frac{\partial p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)/\partial \tau_\nu}{p(b_j|\boldsymbol{w}_K, \boldsymbol{\tau}_K)} \right\} \tag{A.2.17}$$

where, in order to aid readability, the arguments that depend on neither the weight nor lifetime parameters $\boldsymbol{w}_K$ and $\boldsymbol{\tau}_K$ have been dropped. Turning attention now to the derivatives of the bin likelihood, as given by,

$$\begin{aligned} & p(b|\mathcal{H}_K, \boldsymbol{w}_K, \boldsymbol{\tau}_K, \mathcal{I}) \\ & \quad = \frac{|b|}{T} \left(1 - \sum_k w_k \right) + \sum_k w_k \, F(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T_m, \mathcal{I}) \\ & \quad = \frac{|b|}{T} \left(1 - \sum_k w_k \right) + \sum_k w_k \frac{1}{\Lambda} \sum_i \tilde{\gamma}_i \sum_{\ell \geq 0} \Psi_i(\tau_k, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) \end{aligned}$$

and developed in Appendix A.1. Different approximations of the normalising factor $\Lambda(\boldsymbol{w}_K, \boldsymbol{\tau}_K)$ were introduced in Section A.1.5 and, of course, the derivatives required for the Hessian will differ depending on which approximation is used. The first derivatives of the bin-likelihood are easily obtained and are given by,

$$\frac{\partial p(b)}{\partial w_\mu} = \frac{1}{\Lambda}\sum_i \tilde{\gamma}_i \sum_{\ell\geq 0} \Psi_i(\tau_\mu, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) - \frac{|b|}{T} - \frac{\partial \Lambda}{\partial w_\mu}\frac{1}{\Lambda}\left(p(b) - \frac{|b|}{T}\left(1 - \sum_k w_k\right)\right) \tag{A.2.18}$$

$$\frac{\partial p(b)}{\partial \tau_\mu} = \frac{1}{\Lambda}\, w_\mu \sum_i \tilde{\gamma}_i \sum_{\ell\geq 0} \frac{\partial}{\partial \tau_\mu}\Psi_i(\tau_\mu, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) - \frac{\partial \Lambda}{\partial \tau_\mu}\frac{1}{\Lambda}\left(p(b) - \frac{|b|}{T}\left(1 - \sum_k w_k\right)\right) \tag{A.2.19}$$

The second derivatives of the bin-likelihood, again for arbitrary normalising factor Λ (which may or may not depend on the decay signal weight(s) and lifetime(s)), are as follows,

$$\begin{aligned}
&\frac{\partial^2 p(b)}{\partial w_\nu \partial w_\mu} \\
&= \frac{\partial}{\partial w_\nu}\left\{\frac{1}{\Lambda}\left[\sum_i \tilde{\gamma}_i \sum_{\ell\geq 0} \Psi_i(\tau_\mu, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) - \frac{\partial \Lambda}{\partial w_\mu}\left(p(b) - \frac{|b|}{T}\left(1 - \sum_k w_k\right)\right)\right]\right\} \\
&= -\frac{1}{\Lambda}\left[\frac{\partial \Lambda}{\partial w_\nu}\left(\frac{\partial p(b)}{\partial w_\mu} + \frac{|b|}{T}\right) + \frac{\partial \Lambda}{\partial w_\mu}\left(\frac{\partial p(b)}{\partial w_\nu} + \frac{|b|}{T}\right) + \frac{\partial^2 \Lambda}{\partial w_\nu \partial w_\mu}\left(p(b) - w_0\frac{|b|}{T}\right)\right]
\end{aligned} \tag{A.2.20}$$

Using the lifetime derivative (Eqn. A.2.19) as a starting point, the mixed weight-lifetime derivative is given by,

$$
\begin{aligned}
&\frac{\partial^2 p(b)}{\partial w_\nu \partial \tau_\mu} \\
&\quad = \frac{\partial}{\partial w_\nu}\left\{\frac{1}{\Lambda}\left[w_\mu \sum_i \tilde{\gamma}_i \sum_{\ell\geq 0}\frac{\partial}{\partial \tau_\mu}\Psi_i(\tau_\mu, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i)\right.\right. \\
&\qquad \left.\left. -\frac{\partial \Lambda}{\partial \tau_\mu}\left(p(b) - \frac{|b|}{T}\left(1 - \sum_k w_k\right)\right)\right]\right\} \\
&\quad = \delta_{\mu,\nu}\frac{1}{\Lambda}\sum_i \tilde{\gamma}_i \sum_{\ell\geq 0}\frac{\partial}{\partial \tau_\nu}\Psi_i(\tau_\nu, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) \\
&\qquad -\frac{1}{\Lambda}\left[\frac{\partial \Lambda}{\partial w_\nu}\frac{\partial p(b)}{\partial \tau_\mu} + \frac{\partial \Lambda}{\partial \tau_\mu}\left(\frac{\partial p(b)}{\partial w_\nu} + \frac{|b|}{T}\right) + \frac{\partial^2 \Lambda}{\partial w_\nu \partial \tau_\mu}\left(p(b) - w_0\frac{|b|}{T}\right)\right]
\end{aligned}
\tag{A.2.21}
$$

The lifetime only second derivatives are given by,

$$
\begin{aligned}
&\frac{\partial^2 p(b)}{\partial \tau_\nu \partial \tau_\mu} \\
&\quad = \frac{\partial}{\partial \tau_\nu}\left\{\frac{1}{\Lambda}\left[w_\mu \sum_i \tilde{\gamma}_i \sum_{\ell\geq 0}\frac{\partial}{\partial \tau_\mu}\Psi_i(\tau_\mu, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i)\right.\right. \\
&\qquad \left.\left. -\frac{\partial \Lambda}{\partial \tau_\mu}\left(p(b) - \frac{|b|}{T}\left(1 - \sum_k w_k\right)\right)\right]\right\} \\
&\quad = \delta_{\mu,\nu}\frac{1}{\Lambda}\, w_\nu \sum_i \tilde{\gamma}_i \sum_{\ell\geq 0}\frac{\partial^2}{\partial \tau_\nu^2}\Psi_i(\tau_\nu, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) \\
&\qquad -\frac{1}{\Lambda}\left[\frac{\partial \Lambda}{\partial \tau_\mu}\frac{\partial p(b)}{\partial \tau_\nu} + \frac{\partial \Lambda}{\partial \tau_\nu}\frac{\partial p}{\partial \tau_\mu} + \frac{\partial^2 \Lambda}{\partial \tau_\nu \partial \tau_\mu}\left(p(b) - w_0\frac{|b|}{T}\right)\right]
\end{aligned}
\tag{A.2.22}
$$

To proceed it is necessary to determine the (lifetime) derivatives of the quantity $\Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i)$, as defined by Eqn. (A.1.21). Recalling that,

$$\begin{aligned}
&\Psi_i(\tau, b^{\mathrm{L}}, b^{\mathrm{H}}, T, T_m, \sigma_i, \delta_i, u_i) = \\
&\quad \theta[\ell T_m + b^{\mathrm{L}} - \delta_i] \left\{\chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, b^{\mathrm{L}}, \tau, T_m, \sigma_i, \delta_i, u_i)\right\} \\
&\quad + \theta[\ell T_m + b^{\mathrm{H}} - \delta_i] \left\{\chi(\ell, b^{\mathrm{H}}, \tau, T_m, \sigma_i, \delta_i, u_i) - \chi(\ell, \delta_i, \tau, T_m, \sigma_i, \delta_i, u_i)\right\}
\end{aligned}$$

it is clear that this requires the first and second derivatives of the quantity $\chi(\ell, t, \tau, T_m, \sigma, \delta, u)$ to be determined. Recalling that (Eqn. (A.1.20)),

$$\begin{aligned}
&\chi(\ell, t, \tau, T_m, \sigma, \delta, u) \\
&\quad = e^{-(\ell T_m + t - u)/\tau + \frac{1}{2}\sigma^2/\tau^2} \left[\mathrm{erf}\left(\frac{(u - \ell T_m - t)\tau + \sigma^2}{\sigma\tau\sqrt{2}}\right) - \mathrm{erf}\left(\frac{(u - \delta)\tau + \sigma^2}{\sigma\tau\sqrt{2}}\right)\right] \\
&\qquad - \mathrm{erf}\left(\frac{u - \ell T_m - t}{\sigma\sqrt{2}}\right)
\end{aligned}$$

and denoting by $\psi = \psi(\tau)$ the exponent, and by $\theta = \theta(\tau)$ and $\phi = \phi(\tau)$ the arguments to the error functions, such that,

$$\begin{aligned}
\psi &= (u - \ell T_m - t)/\tau + \frac{1}{2}\,\sigma^2/\tau^2, \\
\theta &= \frac{1}{\sigma\sqrt{2}}\left(u - \ell T_m - t + \frac{\sigma^2}{\tau}\right), \\
\phi &= \frac{1}{\sigma\sqrt{2}}\left(u - \delta + \frac{\sigma^2}{\tau}\right)
\end{aligned}$$

the first derivative is determined to be,

$$\frac{\partial\chi}{\partial\tau} = \frac{\partial}{\partial\tau}\left\{e^{\psi}\left[\mathrm{erf}\,(\theta) - \mathrm{erf}\,(\phi)\right]\right\} = \frac{\partial\psi}{\partial\tau}\chi + \frac{2}{\sqrt{\pi}}\frac{\partial\theta}{\partial\tau}\left[e^{\psi-\theta^2} - e^{\psi-\phi^2}\right].$$

The required derivatives are easily determined to be,

$$\frac{\partial\psi}{\partial\tau} = -\frac{1}{\tau^2}\left(u - \ell T_m - t + \sigma^2/\tau\right), \qquad \frac{\partial\theta}{\partial\tau} = \frac{\partial\phi}{\partial\tau} = -\frac{1}{\tau^2}\,\frac{\sigma}{\sqrt{2}}.$$

The second derivative of $\chi(t,\tau)$ is determined in terms of the the variables ψ, θ and ϕ, and their derivatives, and is given by,

$$\begin{aligned}
\frac{\partial^2\chi}{\partial\tau^2} &= \frac{\partial}{\partial\tau}\left\{\frac{\partial\psi}{\partial\tau}\chi + \frac{2}{\sqrt{\pi}}\frac{\partial\theta}{\partial\tau}\left[e^{\psi-\theta^2} - e^{\psi-\phi^2}\right]\right\} \\
&= \frac{\partial^2\psi}{\partial\tau^2}\chi + \frac{\partial\psi}{\partial\tau}\frac{\partial\chi}{\partial\tau} + \frac{2}{\sqrt{\pi}}\frac{\partial^2\theta}{\partial\tau^2}\left[e^{\psi-\theta^2} - e^{\psi-\phi^2}\right] \\
&\quad + \frac{2}{\sqrt{\pi}}\frac{\partial\theta}{\partial\tau}\left[\left(\frac{\partial\psi}{\partial\tau} - 2\theta\frac{\partial\theta}{\partial\tau}\right)e^{\psi-\theta^2} - \left(\frac{\partial\psi}{\partial\tau} - 2\phi\frac{\partial\phi}{\partial\tau}\right)e^{\psi-\phi^2}\right] \\
&= \frac{\partial^2\psi}{\partial\tau^2}\chi + \frac{\partial\psi}{\partial\tau}\frac{\partial\chi}{\partial\tau} + \frac{2}{\sqrt{\pi}}\left[\left(\frac{\partial^2\theta}{\partial\tau^2} + \frac{\partial\theta}{\partial\tau}\left(\frac{\partial\psi}{\partial\tau} - 2\theta\frac{\partial\theta}{\partial\tau}\right)\right)e^{\psi-\theta^2}\right. \\
&\quad \left. - \left(\frac{\partial^2\theta}{\partial\tau^2} + \frac{\partial\theta}{\partial\tau}\left(\frac{\partial\psi}{\partial\tau} - 2\phi\frac{\partial\theta}{\partial\tau}\right)\right)e^{\psi-\phi^2}\right] \\
&= \left(\frac{\partial^2\psi}{\partial\tau^2} + \left(\frac{\partial\psi}{\partial\tau}\right)^2\right)\chi + \frac{2}{\sqrt{\pi}}\left[\left(\frac{\partial^2\theta}{\partial\tau^2} + 2\frac{\partial\theta}{\partial\tau}\left(\frac{\partial\psi}{\partial\tau} - \theta\frac{\partial\theta}{\partial\tau}\right)\right)e^{\psi-\theta^2}\right. \\
&\quad \left. - \left(\frac{\partial^2\theta}{\partial\tau^2} + 2\frac{\partial\theta}{\partial\tau}\left(\frac{\partial\psi}{\partial\tau} - \phi\frac{\partial\theta}{\partial\tau}\right)\right)e^{\psi-\phi^2}\right]
\end{aligned} \tag{A.2.23}$$

where the second derivatives of the quantities ψ, θ, and ϕ, are given by,

$$\frac{\partial^2\psi}{\partial\tau^2} = \frac{1}{\tau^3}\left(2(u - \ell T_m - t) + \frac{3\sigma^2}{\tau}\right), \qquad \frac{\partial^2\theta}{\partial\tau^2} = \frac{1}{\tau^3}\frac{2\sigma}{\sqrt{2}}.$$

Appendix B

Practical TCSPC FLIM: sample preparation and data acquisition

B.1 Preparation of human epithelial *carcinoma* cells

Human epithelial *carcinoma* cells (A431) stably expressing cdc42-GFP [112] were grown in Dulbecco's modified Eagle's medium containing 10% fetal calf serum. Cells were seeded on a glass coverslip (50000 cells/ml) and 48 hours later were fixed with 4% paraformaldehyde, treated with 1mg/ml NaBH4, then mounted with Mowiol (VWR, UK) containing antifade agent and kept at -20 degC.

B.2 Low photon count data acquisition using *Galileo* microscope

Time domain FLIM was performed with an in-house system, excitation being provided by a supercontinuum 40MHz source (SC450-4, Fianium, UK), pulse width $\sim$4 ps, and photon counting performed using a SPC830 TCSPC board (Becker & Hickl, Berlin, Germany). De-scanned detection was afforded by the use of a fast single-photon response, photomultiplier tube (PMH-100-0, Becker & Hickl, Berlin, Germany) placed behind a pinhole, parfocal with the image plane. The following filters were used in data acquisition; 470 nm excitation filter (Semrock FF01-470/22-25, Laser 2000, UK), 495 nm

dichromatic reflector (Semrock FF 495-Di02-25x36, Laser 2000, UK), and 510 nm emission filter (Semrock FF01-510/20-25, Laser 2000, UK) and an IR cut filter. To avoid pulse pile-up, peak photon counting rates remained well below the maximum counting rate offered by the TCSPC electronics[60], with average photon counting rates of the order $10^4 - 10^5$ photons/second. The photon arrival times, with respect to the approximately 40 MHz repetitive laser pulses, were binned into 256 time windows over a total measurement period of 15 ns. Images were captured with a 0.75 NA objective lens (S Fluor 20x/0.75 air, Nikon, UK) at 256×256 pixels corresponding to 334×334 μm at the sample.

The measured instrument response function (IRF) for the Galileo microscope is shown in Fig. B.1, along withe the Bayesian-determined optimal single Gaussian approximation. The portion of the measured instrument response that was used to obtain the results presented in Part III for ML and LS analyses lies inside the vertical lines. It can be seen that the Bayesian single Gaussian approximation and the measured IRF are in close agreement within this region. The effects of detector afterpulsing can be seen in the measured IRF on viewing on a log scale, the ridge to the right of the dominant IRF peak.

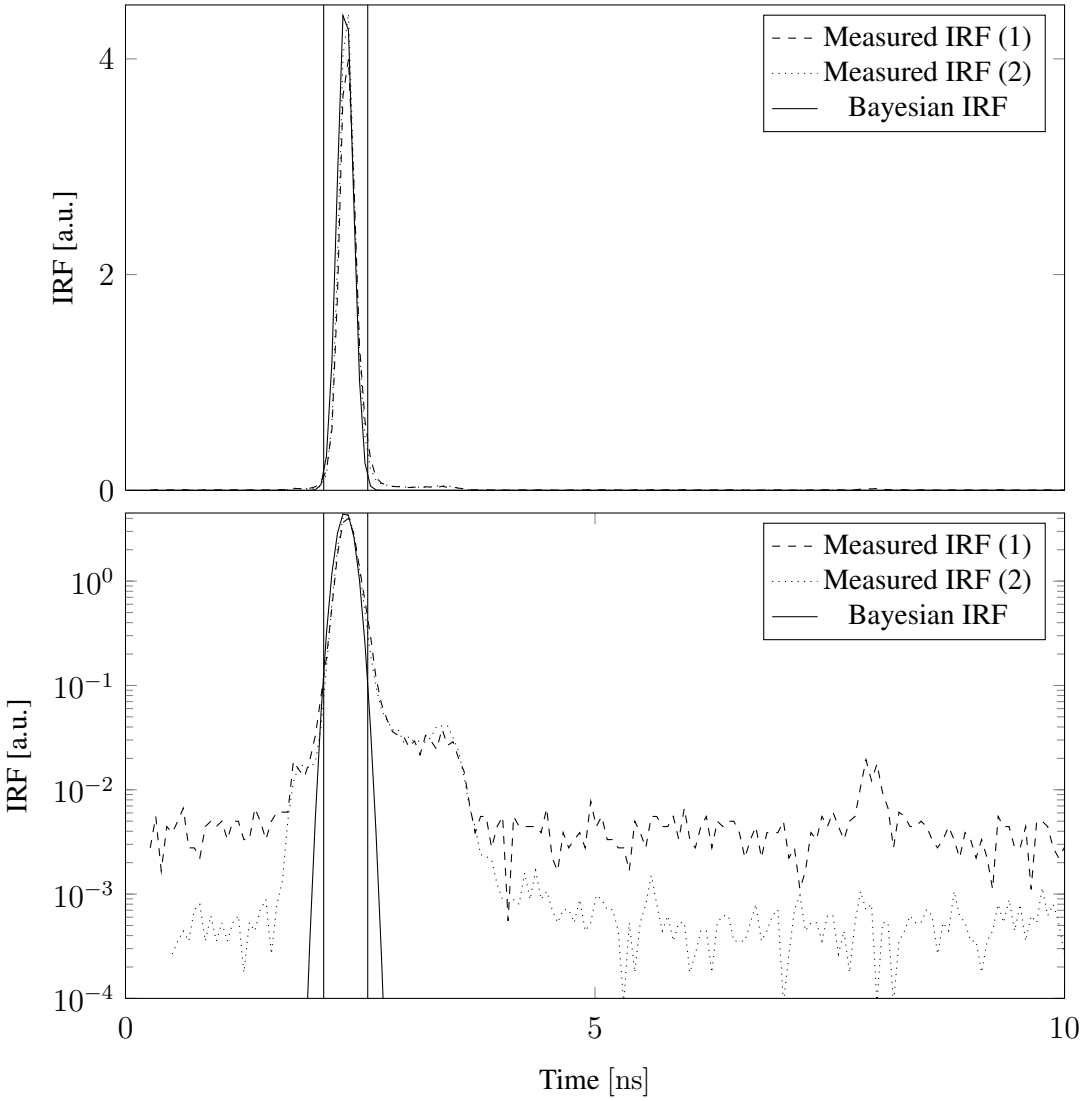


Figure B.1: *Measured IRF and single Gaussian approximation for Galileo microscope:* The measured IRF for the Galileo microscope (Section 1.2.1) and the Bayesian-determined optimal single Gaussian approximation. The The Bayesian-determined optimal single Gaussian instrument response approximation was determined using a single high-count data set (about 35 million photon counts), from the *carcinoma* cell data presented in Section 6.2, resulting from summing the time resolved data from all of the image pixels and on the assumption of mono-exponential decay data; the optimal single Gaussian approximation having been found to have a FWHM width of 0.204 ns (i.e. a standard deviation of 0.086 ns) centered about a delay of 2.341 ns.

Appendix C

Bayesian analysis implementation in C

This chapter provides a summary of the design of the implemented Bayesian analysis algorithms (to realise the Bayesian analysis of Chapter 5), and their incorporation into the existing TRI2 software package [87]. The Bayesian routines have been integrated into TRI2 with the intention of providing an accessible implementation that can be used without possessing any intricate knowledge of either the FLIM system model developed in Chapter 4 or the Bayesian analysis of Chapter 5. At the time of the initial integration of the Bayesian algorithms into TRI2, efforts were made such that their operation should be as similar as possible, from the point of view of user-interaction, to that of the existing LS analysis algorithms, such that a user familiar with the LS routines should be able to use the Bayesian algorithms without facing a steep learning curve.

C.1 Bayesian user-interface in TRI2

The interface to the Bayesian implementation is now introduced, with the aid of screen captures, before discussing some of the technical details of the Bayesian algorithms. On loading a time-resolved image file, and launching the time-resolved analysis panel (and selecting a pixel containing data), a significant amount of information regarding analysis at the selected pixel is presented to the user, as shown in Fig. C.1. The decay parameter estimates resulting from the selected analysis type are shown in the 'Fitted Parameters' panel, the fitted decay is shown in the 'Time Resolved Data and Fit' panel, and the weighted residuals are shown in the 'Difference Between Data and Fit' panel. In the

'Prompt/Excitation' panel, the loaded instrument response and the approximated instrument response used for the Bayesian analysis are displayed.

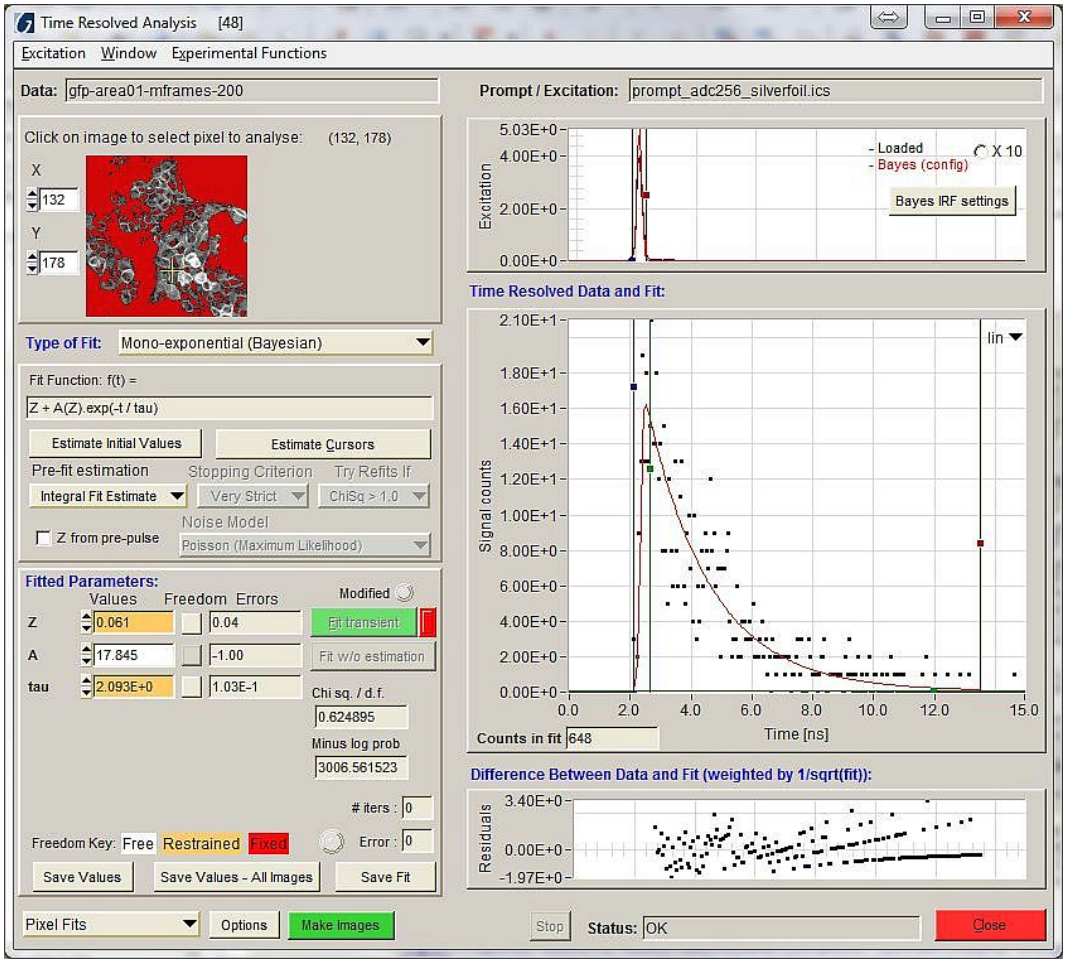


Figure C.1: *Time-resolved fitting panel in TRI2:* The results of mono-exponential Bayesian analysis of an image pixel containing 648 photon counts; notice that the mono-exponential Bayesian parameter estimates and their uncertainty, the fitted fluorescence decay and its residuals, and also the Bayesian approximation to the instrument response and the loaded measured instrument response, are all displayed to the user.

The 'Bayesian Configuration' panel, as shown in Fig. C.2, provides an interface to input user-defined instrument response parameter values (e.g. should the user wish to tweak those values determined from the Bayesian SID analysis (Chapter 8) and also allows the measurement interval and repetition period to be defined (although usually such microscope system configuration details are imported from metadata in the image file).

The ‘Bayesian Configuration’ panel also provides the interface to a rapid Bayesian implementation, the details of which are discussed below.

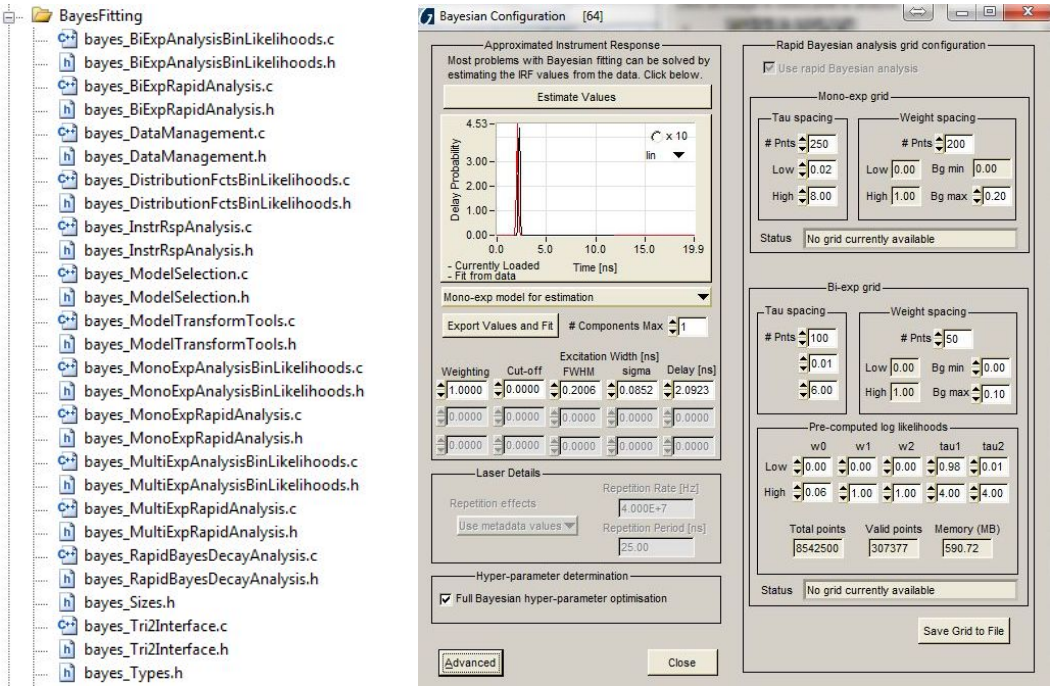


Figure C.2: *Bayesian configuration in TRI2:* This interface stores the Bayesian instrument response approximation parameter values, the repetition period (as accessed in ‘Laser details’), and provides access to the ‘rapid’ Bayesian implementation. To the left of the panel, the ‘tree’ of source files dedicated to the Bayesian implementation is shown.

The ‘Advanced Bayesian Configuration’ panel provides the interface to the optimisation algorithms should the analysis be performed with standard rather than ‘rapid’ algorithms. The panel also enables the Bayesian analysis to be performed with re-binned time-resolved data, as is discussed below. The removal of data that may have been corrupted by consequences of dithering associated with the time-amplitude converter in the TCSPC electronics [60] is also possible from this panel.

Although, of course, the Bayesian algorithms do need to be configuration in order to offer a successful analysis, it is worthwhile noting that (providing TRI2 has been installed with suitable values for the Bayesian configuration parameters) the ‘mechanics’ of the Bayesian analysis are neither presented to the user in the main time-resolved analysis

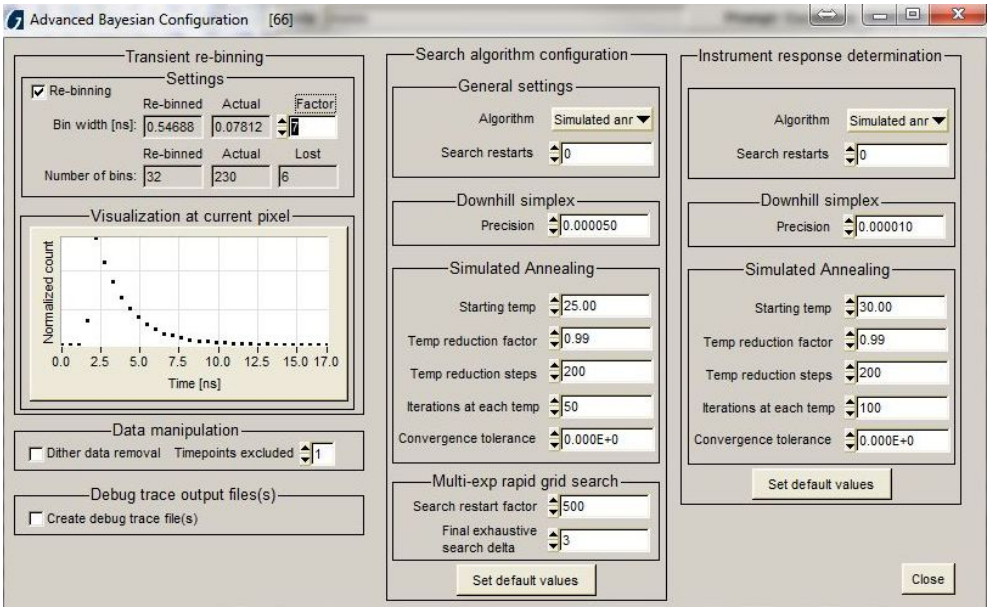


Figure C.3: *Advanced features for Bayesian analysis in TRI2:* Re-binning of transient data is possible from this interface.

panel nor should it usually be necessary for them to be modified.

C.2 Bayesian algorithm design

The Bayesian algorithms have been written using Visual Studio Express 2008 (Microsoft Inc.) as the main development environment, and rely on National Instruments Labwindows/CVI libraries and user-interface components for integration in TRI2. The Bayesian analysis implementation is spread over more than thirty files, as much as possible of each of the different functionalities (i.e. mono-exponential analysis, bi-exponential analysis, Bayesian SID, decay model selection) are separated. All of the Bayesian algorithms have been written with readability, and re-use and portability where appropriate, in mind. Routines that implement variants of an algorithm (e.g. standard mono-exponential analysis and a faster mono-exponential method) have been written with a common interface such that switching between variants merely requires setting the appropriate function pointer. Similarly, all routines that search the parameter space for an optimal value, simulated annealing and downhill simplex routines [113] in the current implementation, have been

developed with identical interfaces, not only for the provision of compact code but also for easy incorporation of additional methods if desired.

The routines have been written to offer robust analysis, and therefore validity checks are made at numerous different points within an analysis. Wherever possible, on detection of an analysis failure, an error code is passed back to the user interface to be translated to meaningful text strings for interpretation by the user.

The interface between the Bayesian algorithms and TRI2 is confined to the files `bayes_Tri2Interface.c` and `bayes_Tri2Interface.h`. All user-interface values or settings that are required by the Bayesian routines are provided through generic wrappers; for example, the following wrapper function,

```
int bayes_GetConfigParameterValueIrComponentsMax(void)
{
    int val;
    GetCtrlVal(BayesConfigPanel,BAYES_IR_COMPONENTS_MAX,&val);
    return (val);
}
```

calls the TRI2 specific routine to obtain the requested value, and then provides it to the Bayesian routine. In implementing the interface to the Bayesian routines in such a way, the Bayesian routines can relatively easily be integrated into a different software package if desired.

A significant number of configuration parameters (e.g. the instrument response approximation parameters, the measurement interval and repetition period, the number of data bins and the data vector) are required by the Bayesian algorithms, although not each routine requires 'sight' of every variable; such parameters populate structures and pointers are passed between the Bayesian routines in order that the number of arguments to a routine is manageable and that the addition of new parameters requires only minor code changes.

C.3 A rapid Bayesian analysis implementation

Two different modes of analysis have been implemented for decay parameter analysis; 'standard' and a 'rapid' parameter estimation. The standard analysis for mono- and bi-

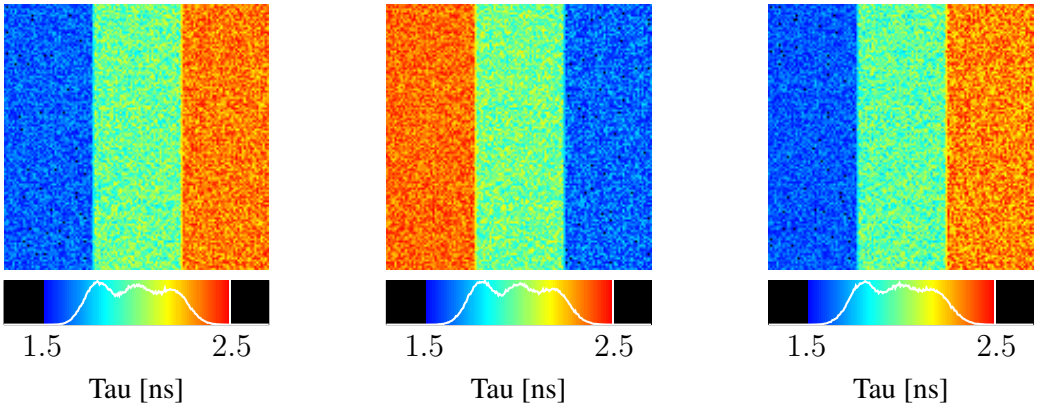
exponential parameter estimation uses the downhill simplex and simulated annealing minimisation algorithms [113], the decay parameters taking values in a 'continuous' parameter space. The 'rapid' analysis has been provided in order to reduce the time taken for image analysis, and decay parameters are estimated using a discrete grid that samples the continuous space.

In adopting such an approach, it is necessary to define the boundaries of the grid; for mono-exponential analysis the boundaries are the minimum and maximum allowed lifetimes, and the minimum and maximum background values; for bi-exponential analysis the grid boundaries are defined by the maximum and minimum values of the two decay weights and lifetimes (as can be seen in Fig. C.2). The fluorescence decay bin-likelihoods need only be determined for those lifetimes on the 'grid', rather than for values from the continuous parameter space as selected by the downhill simplex or simulated annealing algorithms in the 'standard' implementation.

The optimal mono-exponential decay parameters are determined by an exhaustive search of the grid. The implementation of 'rapid' mono-exponential analysis provides a particularly rapid solution as, for reasonable grid sizes, the bin-likelihoods can be pre-computed for each grid point; subsequent analysis of a transient requires only one multiplication and one addition for each bin. Unfortunately, for bi-exponential analysis, it is not possible to pre-compute likelihoods for each point on a grid sufficiently expansive for a reliable analysis; even with only four dimensions (a lifetime and weight for both of the decays) the number of grid points becomes so large that the storage of such a grid is not easily realised.

C.4 Bayesian analysis and transient re-binning

In Section C.1 the option for transient re-binning through the 'Advanced Bayesian Configuration' panel was highlighted. As a means of performing Bayesian analysis more quickly, transient re-binning, has been implemented such that, according to user preference, photon counts in adjacent bins are accumulated into a larger bin (i.e. transient data in 256 bins could be re-binned to be in 16 bins, each spanning 16 of the original 256 bins). Applied to data from Chapter 6, Bayesian analysis of re-binned data is obtained more quickly, and with little change in accuracy even with only 4 bins, as shown in Fig. C.4.



(a) 256 bins (b) 8 bins (c) 4 bins

Figure C.4: *Robust analysis with transient rebinning:* Bayesian-estimated lifetime images for the analysis of synthetic images generated with decay lifetime of 1.8 ns (left panel), 2.0 ns (central panel), and 2.2 ns (right panel), for the different re-binning schemes. In (a) the data (as generated) with 256 time bins, in (b) the same data re-binned into 8 bins, and (c) the same data re-binned into 4 bins. All of the analysed transients were generated to have an intensity of about 1000 total photon counts.

Bibliography

[1] Mark I. Rowley, Paul R. Barber, Anthony C. C. Coolen, and Borivoj Vojnovic, "Bayesian analysis of fluorescence lifetime imaging data," *Proc. SPIE*, vol. 7903, pp. 790325 1–12, 2011.

[2] Tony Ng, Anthony Squire, Gurdip Hansra, Frederic Bornancin, Corinne Prevostel, Andrew Hanby, William Harris, Diana Barnes, Sandra Schmidt, Harry Mellor, Philippe I. H. Bastiaens, and Peter J. Parker, "Imaging Protein Kinase Cα Activation in Cells," *Science*, vol. 283, pp. 2085–2089, 1999.

[3] Fred S. Wouters and Philippe I.H. Bastiaens, "Fluorescence lifetime imaging of receptor tyrosine kinase activity in cells," *Current Biology*, vol. 9, pp. 1127–1130, 1999.

[4] Tuqiang Ni and Lynn A. Melton, "Two-dimensional gas-phase temperature measurements using fluorescence lifetime imaging," *Appl. Spectrosc.*, vol. 50, pp. 1112–1116, 1996.

[5] Richard K. P. Benninger, Oliver Hofmann, James McGinty, Jose Requejo-Isidro, Ian Munro, Mark A. A. Neil, Andrew J. deMello, and Paul M. W. French, "Time-resolved fluorescence imaging of solvent interactions in microfluidic devices," *Optics Express*, vol. 13, pp. 6275–6285, 2005.

[6] Carolyn Jones and Klaus Suhling, "Refractive index sensing using fluorescence lifetime imaging (FLIM)," *Journal of Physics*, vol. 45, pp. 223230, 2006.

[7] James A Levitt, Daniel R Matthews, Simon M Ameer-Beg, and Klaus Suhling, "Fluorescence lifetime and polarization-resolved imaging in cell biology," *Current Opinion in Biotechnology*, vol. 20, pp. 28–36, 2009.

[8] Damian K. Bird, Kent M. Agg, Neil W. Barnett, and Trevor A. Smith, "Time-resolved fluorescence microscopy of gunshot residue: an application to forensic science," *Journal of Microscopy*, vol. 226, pp. 1825, 2007.

[9] Daniela Comelli, Cosimo DAndrea, Gianluca Valentini, Rinaldo Cubeddu, Chiara Colombo, and Lucia Toniolo, "Fluorescence lifetime imaging and spectroscopy as tools for nondestructive analysis of works of art," *Applied Optics*, vol. 43, pp. 2175–2183, 2004.

[10] Daniela Comelli, Gianluca Valentini, Rinaldo Cubeddu, and Lucia Toniolo, "Fluorescence lifetime imaging and fourier transform infrared spectroscopy of Michelangelos David," *Applied Spectroscopy*, vol. 59, pp. 1174–1181, 2005.

[11] Daniela Ribeiro, Jan Willem Borst, Rob Goldbach, and Richard Kormelink, "Tomato spotted wilt virus nucleocapsid protein interacts with both viral glycoproteins gn and gc in planta," *Virology*, vol. 383, pp. 121–130, 2009.

[12] Marjolein Snippea, Jan Willem Borstb, Rob Goldbacha, and Richard Kormelink, "The use of fluorescence microscopy to visualise homotypic interactions of tomato spotted wilt virus nucleocapsid protein in living cells," *Journal of Virological Methods*, vol. 125, pp. 15–22, 2005.

[13] Ionita C. Ghiran, *Light Microscopy: Methods in Molecular Biology*, chapter "Introduction to Fluorescence Microscopy", pp. 93–136, Springer Science + Business Media, 2011.

[14] G. G. Stokes, ""On the Change of Refrangibility of Light"," *Philosophical Transactions of the Royal Society of London*, vol. 142, pp. 463–562, 1852.

[15] J.R. Lakowicz, *Principles of Fluorescence Spectroscopy*, Plenum Press, New York., 1983.

[16] Aleksander Jablonski, "Decay of Photoluminescence of Solutions," *Acta Physica Polonica*, vol. 16, pp. 471–479, 1957.

[17] Aleksander Jablonski, *Luminescence of Organic and Inorganic Materials*, chapter "Decay and Polarization of Fluorescence of Solutions", pp. 110–114, John Wiley and Sons Inc., New York, 1962.

[18] Theodor Förster, "Zwischenmolekulare energiewanderung und fluoreszenz [intermolecular energy migration and fluorescence] (in German)," *Annalen der Physik*, vol. 437, pp. 55–75, 1948.

[19] Theodor Förster, *Modern Quantum Chemistry. Istanbul Lectures. Part III: Action of Light and Organic Crystals*, chapter "Delocalized Excitation and Excitation Transfer", pp. 93–137, New York and London: Academic Press., 1965.

[20] Robert M Clegg, "Fluorescence resonance energy transfer," *Current Opinion in Biotechnology*, vol. 6, pp. 103–110, 1995.

[21] Lubert Stryer, "Fluorescence energy transfer as a spectroscopic ruler," *Ann. Rev. Biochem.*, vol. 47, pp. 819–846, 1978.

[22] Albert H. Coons, Hugh J. Creech, and R. Norman Jones, "Immunological properties of an antibody containing a fluorescent group," *Proc Soc Exp Biol Med*, vol. 47, pp. 200–202, 1941.

[23] Ian D. Odell and Deborah Cook, "Immunofluorescence techniques," *Journal of Investigative Dermatology*, vol. 133, pp. 1–4, 2013.

[24] Jennifer Lippincott-Schwartz and George H. Patterson, "Development and Use of Fluorescent Protein Markers in Living Cells," *Science*, vol. 300, pp. 87–91, 2003.

[25] Elke Haustein and Petra Schwille, "Trends in fluorescence imaging and related techniques to unravel biological information," *HFSP Journal*, vol. 1, pp. 169–180, 2007.

[26] O Shimomura, F H Johnson, and Y Saiga, "Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea," *Journal of cellular and comparitive physiology*, vol. 59, pp. 223–239, 1962.

[27] Roger Y. Tsien, "The green fluorescent protein," *Annu. Rev. Biochem.*, vol. 67, pp. 509–544, 1998.

[28] Ingrid Moen, Charlotte Jevne, Jian Wang, Karl-Henning Kalland, Martha Chekenya, Lars A Akslen, Linda Sleire, Per Ø Enger, Rolf K Reed, Anne M Øyan,

and Linda EB Stuhr, "Gene expression in tumor cells and stroma in dsred 4t1 tumors in egfp-expressing mice with and without enhanced oxygenation," *BMC Cancer*, vol. 12, pp. 21–30, 2012.

[29] Dmitriy M. Chudakov, Sergey Lukyanov, and Konstantin A. Lukyanov, "Fluorescent proteins as a toolkit for *in vivo* imaging," *TRENDS in Biotechnology*, vol. 23, pp. 605–613, 2005.

[30] Ben N. G. Giepmans, Stephen R. Adams, Mark H. Ellisman, and Roger Y. Tsien, "The Fluorescent Toolbox for Assessing Protein Location and Function," *Science*, vol. 312, pp. 217–224, 2006.

[31] Fred S. Wouters, "The physics and biology of fluorescence microscopy in the life sciences," *Contemporary Physics*, vol. 47, pp. 239–255, 2007.

[32] Paul R. Selvin, "The renaissance of fluorescence resonance energy transfer," *nature structural biology*, vol. 7, pp. 730–734, 2000.

[33] Sylvie Lalonde, David W Ehrhardt, and Wolf B Frommer, "Shining light on signaling and metabolic networks by genetically encoded biosensors," *Current Opinion in Plant Biology*, vol. 8, pp. 574–581, 2005.

[34] Wolf B. Frommer, Michael W. Davidson, and Robert E. Campbell, "Genetically encoded biosensors based on engineered fluorescent proteins," *Chemical Society Reviews*, vol. 38, pp. 2833–2841, 2009.

[35] Nicholas Billinton and Andrew W. Knight, "Seeing the Wood through the Trees: A Review of Techniques for Distinguishing Green Fluorescent Protein from Endogenous Autofluorescence," *Analytical Biochemistry*, vol. 291, pp. 175–197, 2001.

[36] Monica Monici, "Cell and tissue autofluorescence research and diagnostic applications," *Biotechnology Annual Review*, vol. 11, pp. 227–256, 2005.

[37] Esther G. Gerrits, Andries J. Smit, and Henk J. G. Bilo, "Ages, autofluorescence and renal function," *Nephrol Dial Transplant*, vol. 24, pp. 710–713, 2009.

[38] Michael Fritzsche and Carl-Fredrik Mandenius, "Fluorescent cell-based sensing approaches for toxicity testing," *Anal Bioanal Chem*, vol. 398, pp. 181–191, 2010.

[39] J.B. Pawley, *The Handbook of Biological Confocal Microscopy*, chapter Fundamental Limits in Confocal Microscopy, pp. 19–37, Plenum Press, New York, 1995.

[40] Nick S. White and Rachel J. Errington, "Fluorescence techniques for drug delivery research: theory and practice," *Advanced Drug Delivery Reviews*, vol. 57, pp. 17–42, 2005.

[41] Thomas H. Chia and Michael J. Levene, "Detection of counterfeit U.S. paper money using intrinsic fluorescence lifetime," *Optics Express*, vol. 17, pp. 22054–22061, 2009.

[42] Klaus Suhling, Nicholas I. Cade, James A. Levitt, Marina K. Kuimova, Pei hua Chung, Gokhan Yahioglu, Gilbert Fruhwirth, Tony Ng, and David Richards, *Biomedical Imaging: The Chemistry of Labels, Probes and Contrast Agents*, chapter Fluorescence Lifetime Imaging applied to Microviscosity Mapping and Fluorescence Modification Studies in Cells, pp. 371–390, Royal Society of Chemistry, 2012.

[43] Klaus Suhling, Paul M. W. French, and David Phillips, "Time-resolved fluorescence microscopy," *Photochemical & Photobiological Sciences*, vol. 4, pp. 13–22, 2005.

[44] Yuansheng Sun, Richard N Day, and Ammasi Periasamy, "Investigating protein-protein interactions in living cells using fluorescence lifetime imaging microscopy," *Nature Protocols*, vol. 6, pp. 1324–1340, 2011.

[45] W. Becker, "Fluorescence lifetime imaging techniques and applications," *Journal of Microscopy*, vol. 247, pp. 119–136, 2012.

[46] D. Phillips, R. C. Drake, D. V. O'Connor, and R. L. Christensen, "Time correlated single-photon counting (TCSPC) using laser excitation," *Instrumentation Science & Technology*, vol. 14, pp. 267–292, 1985.

[47] Jr. Theodorus W.J. Gadella, Thomas M. Jovin, and Robert M . Clegg, "Fluorescence lifetime imaging microscopy (FLIM): Spatial resolution of microstructures on the nanosecond time scale," *Biophysical Chemistry*, vol. 48, pp. 221–239, 1993.

[48] Philippe I. H. Bastiaens and Anthony Squire, "Fluorescence lifetime imaging microscopy: spatial resolution of biochemical processes in the cell," *trends in CELL BIOLOGY*, vol. 9, pp. 48–52, 1999.

[49] Pengguang Wu and Ludwig Brand, "Resonance Energy Transfer: Methods and Applications," *Analytical Biochemistry*, vol. 218, pp. 1–13, 1994.

[50] Ammasi Periasamy, "Fluorescence resonance energy transfer microscopy: a mini review," *Journal of Biomedical Optics*, vol. 6, pp. 287–291, 2001.

[51] Jennifer Lippincott-Schwartz, Erik Snapp, and Anne Kenworthy, "Studying Protein Dynamics in Living Cells," *Molecular Cell Biology*, vol. 2, pp. 444–466, 2001.

[52] Paul J. Tadrous, "Methods for imaging the structure and function of living tissues and cells: 2. Fluorescence lifetime imaging," *Journal of Pathology*, vol. 191, pp. 229–234, 2000.

[53] Ye Chen and Ammasi Periasamy, "Characterization of Two-photon Excitation Fluorescence Lifetime Imaging Microscopy for Protein Localization," *Microscopy Research And Technique*, vol. 63, pp. 72–80, 2004.

[54] W. Becker, A. Bergmann, E. Haustein, Z. Petrasek, P. Schwille, C. Biskup, L. Kelbauskas, K. Benndorf, N. Klöcker, T. Anhut, I. Riemann, and K. König, "Fluorescence Lifetime Images and Correlation Spectra Obtained by Multidimensional Time-Correlated Single Photon Counting," *Microscopy Research And Technique*, vol. 69, pp. 186–195, 2006.

[55] Desmond V O'Connor and David Phillips, *Time-correlated Single Photon Counting*, ACADEMIC PRESS INC. (LONDON) LTD, 1984.

[56] C. J. de Grauw and H. C. Gerritsen, "Multiple time-gate module for fluorescence lifetime imaging," *Applied Spectroscopy*, vol. 55, pp. 670–678, 2001.

[57] H. C. Gerritsen, M. A. H. Asselbergs, A. V. Agronskaia, and W. G. J. H. M. Van Sark, "Fluorescence lifetime imaging in scanning microscopes: acquisition speed, photon economy and lifetime resolution," *Journal of Microscopy*, vol. 206, pp. 218–224, 2002.

[58] Thomas M. Nordlund, *Topics in Fluorescence Spectroscopy*, chapter Streak Cameras for Time-Domain Fluorescence, pp. 183–260, Springer US, 2002.

[59] R. V. Krishnan, H. Saitoh, H. Terada, V. E. Centonze, and B. Herman, "Development of a multiphoton fluorescence lifetime imaging microscopy system using a streak camera," *Review Of Scientific Instruments*, vol. 74, pp. 2714–2721, 2003.

[60] Wolfgang Becker, *The bh TCSPC Handbook*, Becker & Hickl GmbH, Third edition, 2008.

[61] J. Requejo-Isidro, J. McGinty, I. Munro, D. S. Elson, N. P. Galletly, M. J. Lever, M. A. A. Neil, G. W. H. Stamp, P. M. W. French, P. A. Kellett, J. D. Hares, and A. K. L. Dymoke-Bradshaw, "High-speed wide-field time-gated endoscopic fluorescence-lifetime imaging," *Optics Letters*, vol. 29, pp. 2249–2251, 2004.

[62] P R Barber, I D C Tullis, G P Pierce, R G Newman, J Prentice, M Rowley, D Matthews, S M Ameer-Beg, and B Vojnovic, "The Gray Institute Open High-Content, Fluorescence Lifetime Microscopes," J Microscopy (submitted December 2012).

[63] Bryan Q. Spring and Robert M. Clegg, *FLIM Microscopy in Biology and Medicine*, chapter Frequency-Domain FLIM, pp. 115–142, Chapman and Hall/CRC, 2009.

[64] Zelkjo Bajzer, Terry M. Therneau, Joseph C. Sharp, and Franklin G. Prendergast, "Maximum likelihood method for the analysis of time-resolved fluorescence decay curves," *European Biophysics Journal*, vol. 20 (5), pp. 247–262, 1991.

[65] Peter J. Verveer, Anthony Squire, and Philippe I. H. Bastiaens, "Global analysis of fluorescence imaging microscopy data," *Biophysical Journal*, vol. 78 (4), pp. 2127–2137, 2000.

[66] S. Pelet, M. J. R. Previte, L. H. Laiho, and P. T. C. So, "A fast global fitting algorithm for fluorescence lifetime imaging microscopy based on image segmentation," *Biophysical Journal*, vol. 84, pp. 2807–2817, 2004.

[67] P. R. Barber, S. M. Ammer-Beg, J. Gilbey, R. J. Edens, I. Ezike, and B. Vojnovic, "Global and pixel kinetic data analysis for FRET detection by multi-photon time-domain FLIM," *Proc. SPIE*, vol. 5700, pp. 171–181, 2005.

[68] Javier A. Jo, Qiyin Fang, and Laura Marcu, "Ultrafast Method for the Analysis of Fluorescence Lifetime Imaging Microscopy Data Based on the Laguerre Expansion Technique," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 11 (4), pp. 835–845, 2005.

[69] Glen I. Redford and Robert M. Clegg, "Polar plot representation for frequency-domain analysis of fluorescence lifetimes," *Journal of Fluorescence*, vol. 15 (5), pp. 805–815, 2005.

[70] Michelle A. Digman, Valeria R. Caiolfa, Moreno Zamai, and Enrico Gratton, "The phasor approach to fluorescence lifetime imaging analysis," *Biophysical Journal*, vol. 94 (2), pp. L14–L16, 2007.

[71] Sergi Padilla-Parra, Nicolas Auduge, Maite Coppey-Moisan, and Marc Tramier, "Quantitive FRET analysis by fast acquisition time domain FLIM at high spatial resolution in living cells," *Biophysical Journal*, vol. 95 (6), pp. 2976–2988, 2008.

[72] K. C. Benny Lee, J. Siegel, S. E. D. Webb, S. Leveque-Fort, M. J. Cole, R. Jones, K. Dowling, M. J. Lever, and P. M. W. French, "Application of the stretched exponential function to fluorescence lifetime imaging," *Biophysical Journal*, vol. 81, pp. 12651274, 2001.

[73] Borys Kierdaszuk, "From discrete multi-exponential model to lifetime distribution model and power law fluorescence decay function," *Spectroscopy*, vol. 24, pp. 399407, 2010.

[74] Goro Nishimura and Mamoru Tamura, "Artefacts in the analysis of temporal response functions measured by photon counting," *Physics in Medicine and Biology*, vol. 50, pp. 1327–1342, 2005.

[75] Peter Hall and Ben Sellinger, "Better estimates of exponential decay parameters," *Journal of Physical Chemistry*, vol. 85 (20), pp. 2941–2946, 1981.

[76] David A. Turton, Gavin D. Reid, and Godfrey S. Beddard, "Accurate analysis of fluorescence decays from single molecules in photon counting experiments," *Analytical Chemistry*, vol. 75 (16), pp. 4182–4187, 2003.

[77] Joel Tellinghuisen, "Bias and precision in the estimation of exponential decay parameters from sparse data," *Analytical Chemistry*, vol. 65 (9), pp. 1240–1246, 1993.

[78] Joseph J. K. Ó Ruanaidh and William J. Fitzgerald, *Numerical Bayesian Methods Applied to Signal Processing*, Springer, 1996.

[79] Emery N. Brown, Zhenhua Zhang, and Alex D. McCollom, "On the uncertainty in single molecule fluorescent lifetime and enery emission measurements," Tech. Rep., United States. National Aeronautics and Space Administration, 1994.

[80] Regina Won Kay Leung, Shu-Chi Allison Yeh, and Qiyin Fang, "Effects of incomplete decay in fluorescence lifetime estimation," *Biomedical Optics Express*, vol. 2, pp. 2517–2531, 2011.

[81] Sang Hoon Lee, Moosung Jae, and Robin P. Gardner, "Non-Poisson counting statistics of a hybrid G-M counter dead time model," *Beam Interactions with Materials & Atoms*, vol. 263, pp. 46–49, 2007.

[82] Ted A Laurence and Brett A Chromy, "Efficient maximum likelihood estimator fitting of histograms," *Nature Methods*, vol. 7, no. 5, pp. 338–339, 2010.

[83] Richard M. Ballew and J. N. Demas, "An error analysis of the rapid lifetime determination method for the evaluation of single exponential decays," *Anal. Chem.*, vol. 61, pp. 30–33, 1989.

[84] Kristin K. Sharman, Ammasi Periasamy, Harry Ashworth, J. N. Demas, and N. H. Snow, "Error analysis of the rapid lifetime determination method for double-exponential decays and new windowing schemes," *Anal. Chem.*, vol. 71, pp. 947–952, 1999.

[85] Joseph R. Lakowicz, Gabor Laczko, Henryk Cherek, Enrico Gratton, and Mark Limkeman, "Analysis of Fluorescence Decay Kinetics from Variable-frequency Phase Shift and Modulation Data," *Biophys. J.*, vol. 46, pp. 463–477, 1984.

[86] P. J. Verveer and P. I. H. Bastiaens, "Evaluation of global analysis algorithms for single frequency fluorescence lifetime imaging microscopy data," *Journal of Microscopy*, vol. 209, pp. 1–7, 2003.

[87] P. R. Barber, S. M. Ammer-Beg, J. Gilbey, L. M. Carlin, M. Keppler, T. C. Ng, and B. Vojnovic, "Multiphoton time-domain fluorescence lifetime imaging microscopy: practical application to protein-protein interactions using global analysis," *Journal of The Royal Society Interface*, vol. 6, pp. S93–S105, 2009.

[88] Hernan E. Grecco, Pedro Roda-Navarro, and Peter J. Verveer, "Global analysis of time correlated single photon counting FRET-FLIM data," *Optics Express*, vol. 17, pp. 6493–6508, 2009.

[89] Alessandro Esposito, Hans C. Gerritsen, and Fred S. Wouters, "Fluorescence lifetime heterogeneity resolution in the frequency domain by lifetime moments analysis," *Biophysical Journal*, vol. 89, pp. 42864299, 2005.

[90] A. H. A. Clayton, Q. S. Hanley, and P. J. Verveer, "Graphical representation and multicomponent analysis of single-frequency fluorescence lifetime imaging microscopy data," *Journal of Microscopy*, vol. 213, pp. 1–5, 2004.

[91] Fred S. Wouters and Alessandro Esposito, "Quantitive analysis of fluorescence lifetime imaging made easy," *HFSP Journal*, vol. 2 (1), pp. 7–11, 2008.

[92] Aditi S. Dabir, Chintan A. Trivedi, Yeontack Ryu, Paritosh Pande, and Javier A. Jo, "Fully automated deconvolution method for on-line analysis of time-resolved fluorescence spectroscopy data based on an iterative laguerre expansion technique," *Journal of Biomedical Optics*, vol. 14 (2), 2009.

[93] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.h. Ackerman, "Exponential parameter estimation (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance Part A*, vol. 27A, pp. 5563, 2005.

[94] G. Larry Bretthorst, William C. Hutton, Joel R. Garbow, and Joseph J.h. Ackerman, "Exponential model selection (in NMR) using Bayesian probability theory," *Concepts in Magnetic Resonance Part A,*, vol. 27A, pp. 6472, 2005.

[95] Paul R Barber, Simon M Ameer-Beg, Senthila Pathmananthan, Mark Rowley, and A C C Coolen, "A Bayesian method for single molecule, fluorescence burst analysis," *Biomedical Optics Express*, vol. 1, no. 4, pp. 1148–1158, 2010.

[96] (ed) Neil D. Lawrence, Mark Girolami, Magnus Rattray, and Guido Sanguinetti, *Learning and Inference in Computational Systems Biology*, The MIT Press, 2010.

[97] Mingjun Zhong, Mark Girolami, Karen Faulds, and Duncan Graham, "Bayesian methods to detect dye-labelled DNA oligonucleotides in multiplexed Raman spectra," *Journal of the Royal Statistical Society*, vol. 60, pp. 187–206, 2011.

[98] A. E. W. Knight and B. K. Selinger, "Single photon decay spectroscopy," *Australian Journal of Chemistry*, vol. 26 (1), pp. 1–27, 1973.

[99] Chris M. Harris and Ben K. Selinger, "Single-photon decay spectroscopy. II, the pile-up problem," *Australian Journal of Chemistry*, vol. 32 (10), pp. 2111–2129, 1979.

[100] Malte Kollner and Jurgen Wolfrum, "How many photons are necessary for fluorescence-lifetime measurements?," *Chemical Physics Letters*, vol. 200 (1–2), pp. 199–204, 1992.

[101] Alessandro Esposito, Hans C. Gerritsen, and Fred S. Wouters, "Optimizing frequency-domain fluorescence lifetime sensing for high-throughput applications: photon economy and acquisition speed," *J. Opt. Soc. Am, A*, vol. 24, pp. 3261–3273, 2007.

[102] Leo M. Carlin, Rachel Evans, Hanna Milewicz, Luis Fernandes, Daniel R. Matthews, Michela Perani, James Levitt, Melanie D. Keppler, James Monypenny, Ton Coolen, Paul R. Barber, Borivoj Vojnovic, Klaus Suhling, Franca Fraternali, Simon Ameer-Beg, Peter J. Parker, N. Shaun B. Thomas, and Tony Ng, "A Targeted siRNA Screen Identifies Regulators of Cdc42 Activity at the Natural Killer Cell Immunological Synapse," *Science Signaling*, vol. 4, pp. ra81, 2011.

[103] G A Morton, H M Smith, and R Wasserman, "Afterpulses in photomultipliers," *IEEE Transactions on Nuclear Science*, vol. 14, pp. 443–448, 1967.

[104] P B Coates, "The origins of afterpulses in photomultipliers," *J. Phys. D: Appl. Phys.*, vol. 6, pp. 1159–1166, 1973.

[105] A. Leray, C. Spriet, D. Trinel, and Laurent Héliot, "Three-Dimensional Polar Representation for Multispectral Fluorescence Lifetime Imaging Microscopy," *Cytometry Part A*, vol. 75A, pp. 1007–1014, 2009.

[106] Mariusz Szabelski, Rafal Luchowski, Zygmunt Gryczynski, Peter Kapusta, Uwe Ortmann, and Ignacy Gryczynski, "Evaluation of instrument response functions for lifetime imaging detectors using quenched rose bengal solutions," *Chemical Physics Letters*, vol. 471, pp. 153–159, 2009.

[107] Raluca Niesner and Karl heinz Gericke, "Fluorescence lifetime imaging in biosciences: technologies and applications," *Front. Phys. China*, vol. 3, pp. 88–104, 2008.

[108] Lin Shi, Wen Liu, Heye Zhang, Yongming Xie, and Defeng Wang, "A survey of GPU-based medical image computing techniquess," *Quantitative Imaging in Medicine and Surgery*, vol. 2, pp. 188–206, 2012.

[109] Yina Wang, Tingwei Quan, Shaoqun Zeng, and Zhen-Li Huang, "PALMER: a method capable of parallel localization of multiple emitters for high-density localization microscopy," *Optics Express*, vol. 20, pp. 16039–16049, 2012.

[110] Christopher M. Bishop, *Neural Networks for Pattern Recognition*, Oxford University Press, 1995.

[111] David J. C. MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003.

[112] Konstantina Makrogianneli, Leo M. Carlin, Melanie D. Keppler, Daniel R. Matthews, Enyinnaya Ofo, Anthony Coolen, Simon M. Ameer-Beg, Paul R. Barber, Borivoj Vojnovic, and Tony Ng, "Integrating Receptor Signal Inputs That Influence Small Rho GTPase Activation Dynamics at the Immunological Synapse," *Molecular and cellular biology*, vol. 29, pp. 2997–3006, 2009.

[113] *Numerical Recipes in C, The Art of Scientific Computing (2nd Edition)*, Cambridge University Press, 1992.